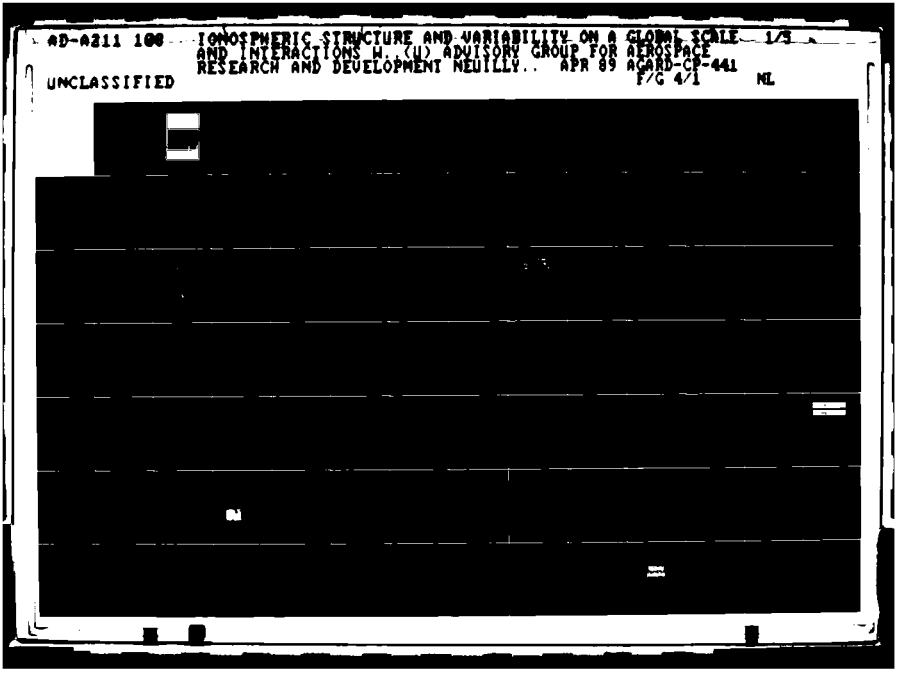
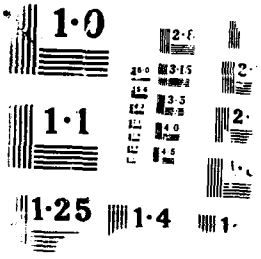
AD-A211 108 IONOSPHERIC STRUCTURE AND VARIABILITY ON A GLOBAL SCALE 1/3
AND INTERACTIONS W..(U) ADVISORY GROUP FOR AEROSPACE
RESEARCH AND DEVELOPMENT NEUILLY.. APR 89 AGARD-CP-441
UNCLASSIFIED F/G 4/1 NL




1·0
1·1
1·25
1·4

AD-A211108

AGARD-CP-441

AGARD-CP-441

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AGARD CONFERENCE PROCEEDINGS No.441

# Ionospheric Structure and Variability on a Global Scale and Interactions with Atmosphere and Magnetosphere

DTIC
ELECTE
JUN 23 1989
S D

NORTH ATLANTIC TREATY ORGANIZATION



DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

DISTRIBUTION AND AVAILABILITY ON BACK COVER

89 6 22 055

AGARD-CP-441

NORTH ATLANTIC TREATY ORGANIZATION

ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT

(ORGANISATION DU TRAITE DE L'ATLANTIQUE NORD)

AGARD Conference Proceedings No.441

# IONOSPHERIC STRUCTURE AND VARIABILITY ON A GLOBAL SCALE AND INTERACTIONS WITH ATMOSPHERE AND MAGNETOSPHERE

Papers presented at the Electromagnetic Wave Propagation Panel Symposium held in Munich, Germany, 16—20 May 1988.

## THE MISSION OF AGARD

According to its Charter, the mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

- Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community;
- Providing scientific and technical advice and assistance to the Military Committee in the field of aerospace research and development (with particular regard to its military application);
- Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;
- Improving the co-operation among member nations in aerospace research and development;
- Exchange of scientific and technical information;
- Providing assistance to member nations for the purpose of increasing their scientific and technical potential;
- Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Programme and the Aerospace Applications Studies Programme. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

The content of this publication has been reproduced
directly from material supplied by AGARD or the authors.

Published April 1989

Copyright © AGARD 1989
All Rights Reserved

ISBN 92-835-0500-X

*Printed by Specialised Printing Services Limited*
*40 Chigwell Lane, Loughton, Essex IG10 3TZ*

## FOREWORD

It is now well-known that the magnetosphere, ionosphere, and thermosphere are strongly coupled via electric fields, particle precipitation, field-aligned currents, heat flows and frictional interactions. It has also been clearly established that the entire system is driven by temporal variations of the solar wind and interplanetary magnetic field (IMF). For example, variations in the solar wind dynamic pressure and the IMF direction affect the magnetospheric electric field and energetic particle precipitation patterns. Changes in these patterns then affect the ionospheric structure, composition, and circulation, and these changes, in turn, affect the thermosphere. During the last decade, significant progress has been made in elucidating the main coupling processes linking the magnetosphere, ionosphere and thermosphere. Associated with this advance has been a parallel advance in our understanding of the formation and evolution of the major morphological features observed in the ionosphere. However, in addition to the large-scale morphological features, the ionosphere also exhibits a considerable amount of structure, with scale lengths ranging from metres to hundreds of kilometres. At present, the formation and evolution of ionospheric structure is less clearly understood.

Because of the strong coupling, the time delays, and the feedback mechanisms that exist between the different elements of the magnetosphere-ionosphere-thermosphere system, a complete understanding of ionospheric behaviour can be achieved only when the entire system is studied. With this in mind, the Electromagnetic Wave Propagation Panel sponsored a symposium on "Ionospheric Structure and Variability on a Global Scale and Interaction with the Atmosphere and Magnetosphere". The symposium was held in Munich, Germany, 16—20 May, 1988. The meeting was divided into six sessions, including Solar-Terrestrial Relationships, Ionosphere-Magnetosphere Interactions, Global Ionospheric Dynamics, Ionosphere-Atmosphere Interactions, Ionosphere Irregularities, and Multistation/Multiparameter Observations and Systems Impact. In each session there were both invited and contributed presentations, with a total of 46 presentations during the five-day meeting. These proceedings contain the papers presented at the symposium; they were reproduced directly from copies furnished by the authors.

We would like to thank the Programme Committee Members and Session Chairmen for their help. Special thanks are also due to the EPP Executive, the German Delegation, the interpreters and the AGARD staff.

L.Bossy
R.W.Schunk




Accesion For
NTIS CRA&I ☒
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution /
Availability Codes
Dist | Avail and / or Special
A-1

## THEME

Hitherto, study of the aerospace propagation media tended to concentrate on localised phenomena or events. However, there is a range of phenomena which need for their elucidation observations and analysis on a global scale. Much of the local ionospheric behaviour is determined by coupling electrically to the distant ionosphere and solar wind, and coupling dynamically and electrodynamically to higher and lower levels of the atmosphere. Global multistation/multiparameter observations are often found to be necessary to provide the frame of reference for interpreting both stationary and propagating locally observed effects.

Communications, Navigation and Surveillance systems operating in/through the aerospace EM propagation environment are affected by the state/variability of the propagation media. An understanding of the complex global interaction would improve the means of predictability and assessment of localised phenomena and suggest methods for mitigation of adverse propagation conditions.

* * *

Jusqu'à présent, l'étude des milieux de propagation aérospatiaux avait tendance à porter essentiellement sur des phénomènes ou des évènements localisés. Il existe cependant toute un série de phénomènes qui, pour être bien compris, nécessitent des observations et une analyse à l'échelle du globe. Le comportement de l'ionosphère locale est déterminé en grand partie par son couplage dynamique et électrodynamique avec les couches inférieure et supérieure de l'atmosphère. Des observations à l'échelle du globe portant sur de nombreux paramètres relevés par un grand nombre de stations s'avèrent souvent nécessaires pour fournir un cadre de référence permettant d'interpréter les effets stationnaires et propagés observés localement.

Les systèmes de communications, de navigation et de surveillance fonctionnant dans ou à travers le milieu aérospatial de propagation électromagnétique sont affectés par l'état ou la variabilité du milieu dans lequel se produit la propagation. La compréhension de l'interaction complexe à l'échelle du globe permettrait d'améliorer les moyens utilisés pour prévoir et évaluer les phénomènes localisés et de proposer des méthodes susceptibles de rendre moins sévères les conditions défavorables de propagation.

## ELECTROMAGNETIC WAVE PROPAGATION PANEL


Chairman: Prof. C.Goutelard
Laboratoire d'Etude de Transmissions Ionosphériques
Université Paris Sud
9, Avenue de la Division Leclerc
94230 Cachan
France

Deputy Chairman: Ir. H.Vissinga
Physics and Electronics Lab.
TNO
P.O. Box 96864
2509 JG The Hague
Netherlands


## TECHNICAL PROGRAMME COMMITTEE

### CO-CHAIRMEN


Prof. L.Bossy
174, Avenue W.Churchill
UCCLE
B-1180 Brussels
Belgium

Prof. R.W.Schunk
Center of Atmospheric and Space Sciences
UMC 4405
Utah State University
Loga, Utah 84322-4405
United States


### COMMITTEE


Dr H.Soicher
Center for Command Control & Comm. Syst.
U.S. Army Communications-Electronics Command
ATTN: AMSEL-RD-C3-TA-1
Fort Monmouth, N.J. 07703-5202
United States

Prof. T.B.Jones
University of Leicester
Department of Physics
University Road
Leicester LE2 7RH
United Kingdom

Dr A.Brekke
The Surdral Observ.
PO Box 953
N 9001 Tromso
Norway

Prof. R.D.Hunsucker
Geophysical Institute
University of Alaska
Fairbanks, AK 99775-0800
United States

Dr K.Schlegel
Max Planck Institut für Aeronomie
*Postfach 20*
D-3411 Katlenburg-Lindau
Federal Republic of Germany

Dr C.Senior
CNET /CRPE
38—40 rue du Général Leclerc
92131 Issy Les Moulineaux
France

Dr G.Rostoker
University of Alberta
Department of Physics
Edmonton, Alberta T6G 2J1
Canada


## ELECTROMAGNETIC WAVE PROPAGATION PANEL EXECUTIVE

LTC P.A.Brunelli


From Europe
NATO-AGARD
7, rue Ancelle
92200, Neuilly-sur-Seine
France

From the USA and Canada
NATO-AGARD
APO New York 09777

CONTENTS


| | Page |
|---|---|
| FOREWORD | iii |
| THEME | iv |
| ELECTROMAGNETIC WAVE PROPAGATION PANEL OFFICERS AND COMMITTEE | v |
| TECHNICAL EVALUATION REPORT<br>by R.W.Schunk | ix |
| | Reference |
| SOLAR-TERRESTRIAL RELATIONSHIP | |
| SOLAR-TERRESTRIAL RELATIONS: FLARE AND SOLAR WIND EFFECTS<br>by J.Lemaire | 1 |
| MAGNETOSPHERIC SUBSTORMS AS A SIGNATURE OF THE SOLAR TERRESTRIAL INTERACTION<br>by G.Rostoker | 2 |
| BALANCE EQUATIONS FOR IONOSPHERIC PLASMAS WITH DIFFERENT PARTIAL TEMPERATURE<br>by K.Suchy | 3 |
| NON-LINEAR WAVE-EQUATIONS FOR LOW-FREQUENCY ACOUSTIC GRAVITY WAVES<br>by R.H.M.Miesen, L.P.J.Kamp and F.W.Sluijter | 4 |
| GLOBAL IONOSPHERIC DYNAMICS | |
| GLOBAL IONOSPHERIC DYNAMICS<br>by K.Rawer | 5 |
| GLOBAL IONOSPHERIC DYNAMICS: A REVIEW<br>by H.Rishbeth | 6 |
| A FULLY ANALYTIC, LOW AND MIDDLE LATITUDE IONOSPHERIC MODEL<br>by D.N.Anderson and J.M.Forbes | 7 |
| GLOBAL MORPHOLOGY OF PLASMA BUBBLES IN THE LOW LATITUDE IONOSPHERE<br>by K.Schlegel, K.Oyama and T.Takahashi | 8 |
| SIMULATION OF THE POLAR CAP F REGION IONIZATION USING AN EXPERIMENTAL CONVECTION ELECTRIC FIELD<br>by C.Taieb | 9* |
| AVERAGE DIURNAL AND SEASONAL PROFILES OF IONOSPHERIC STORMS<br>by G.L.Wrenn, A.S.Rodger and H.Rishbeth | 10 |
| MEDIUM SCALE STRUCTURE OF THE F-REGION<br>by A.K.Paul | 11 |
| VLF/LF RADIO WAVE STUDIES OF THE STRUCTURE AND VARIABILITY OF THE EQUATORIAL, MID-LATITUDE AND POLAR IONOSPHERE BELOW 100 Km<br>by P.A.Kossey, J.E.Rasmussen and W.I.Klemetti | 12 |
| THE MIDDLE AND HIGH LATITUDE IONOSPHERE AT ABOUT 550 Km ALTITUDE<br>by Y.K.Tulunay | 13 |
| MODELLING IONOSPHERIC DENSITY STRUCTURES<br>by R.W.Schunk and J.J.Sojka | 14 |


* Abstract only.

Reference

IONOSPHERE/MAGNETOSPHERE INTERACTIONS

Paper 15 withdrawn

OBSERVATIONS OF IONOSPHERE/MAGNETOSPHERE INTERACTIONS FROM THE DYNAMICS EXPLORER SATELLITES
by R.A.Hoffman 16

EMPIRICAL MODELS OF CONVECTION ELECTRIC FIELDS AND ELECTROSTATIC POTENTIAL AT HIGH LATITUDE FROM EISCAT OBSERVATIONS: PRELIMINARY RESULTS
by C.Senior, D.Fontaine and G.Caudal 17*

Paper 18 withdrawn

LOW FREQUENCY ELECTROSTATIC WAVES OBSERVED IN THE VICINITY OF AN AURORAL ARC
by K.Rinnert 19

POLAR F LAYER "LACUNA" IRREGULARITIES AT DUMONT D'URVILLE
by P.Vila 20

IONOSPHERE/ATMOSPHERE INTERACTIONS

SEASONAL AND GEOMAGNETIC RESPONSE OF THE THERMOSPHERE AND IONOSPHERE
by D.Rees and T.J.Fuller-Rowell 21

STORMTIME COUPLING OF THE NEUTRAL AND IONIZED UPPER ATMOSPHERE AT MIDDLE LATITUDES
by G.W.Prölss 22

IONOSPHERE-THERMOSPHERE COUPLING AT HIGH LATITUDES: ASPECT ANGLE DEPENDENCE OF NON-THERMAL PLASMAS
by K.J.Winser, et al. 23*

QUIET TIME CONDUCTIVITIES OF THE AURORAL IONOSPHERE
by A.Brekke and C.Hall 24

CALCULATION OF THE NEUTRAL GAS TEMPERATURE AND VELOCITY VECTOR OF THE NEUTRAL GAS USING EISCAT CP-3 DATA
by D.Figueroa and H.Kohl 25

ION HEATING EVENTS OBSERVED BY THE EISCAT RADAR
by I.W.McCrea, T.R.Robinson, M.Lester and T.B.Jones 26

ON THE INFLUENCE OF IONOSPHERIC CONDUCTIVITY ON DISPERSION OF ACOUSTIC GRAVITY WAVES
by L.P.J.Kamp, R.H.M.Miesen, P.C. de Jagher and F.W.Sluijter 27

ATMOSPHERIC GRAVITY WAVES IN THE AURORAL SOURCE REGION OF TIDs STUDIED WITH THE EISCAT CP2-DATA
by H.P.Mauelshagen and K.Schlegel 28

ON THE POSSIBILITY OF PRODUCING ARTIFICIAL IONIZATION AND POLAR AURORA IN THE IONOSPHERE AND BY RADIO-WAVES EMITTED FROM THE GROUND
by M.Cutolo and A.Argenziano 29

IONOSPHERIC IRREGULARITIES

IONOSPHERIC IRREGULARITIES DUE TO POWERFUL HF RADIO TRANSMISSIONS
by J.A.Fejer 30

* Abstract only

Reference

A REVIEW ON RADIO STUDIES OF AURORAL E-REGION IONOSPHERIC IRREGULARITIES
by C.Haldoupis 31

IONOSPHERIC IRREGULARITIES AND THEIR EFFECTS ON ELECTROMAGNETIC WAVES PROPAGATING THROUGH THEM
by K.C.Yeh 32

ELECTROSTATIC TURBULENCE IN THE HIGH LATITUDE IONOSPHERE
by J.C.Cerisier, H.Mounir and J.-J.Berthelier 33*

AURORAL AND SUB-AURORAL F-LAYER IRREGULARITY STUDIES IN THE NORTHERN AND SOUTHERN HEMISPHERE DURING THE EQUINOX TRANSITION STUDY, SEPTEMBER 16–24, 1984
by J.Aarons, J.C.Foster, L.Kersley and A.S.Rodger 34

SMALL-SCALE IRREGULARITIES AND GLOBAL PLASMA DYNAMICS IN THE HIGH-LATITUDE IONOSPHERE
by C.Hanuise, et al. 35

EXAMPLES OF METEOROLOGICAL BEHAVIOUR OF THE IONOSPHERE
by K.Bibl 36

MULTISTATION/MULTIPARAMETER OBSERVATIONS AND SYSTEMS IMPACT

MULTISTATION/MULTIPARAMETER OBSERVATIONS WITH A NETWORK OF DIGITAL IONOSONDES
by B.W.Reinisch, J.Buchau, K.Bibl and G.S.Sales 37

COORDINATED MULTI-PARAMETER MEASUREMENTS AND PREDICTIONS OF THE GLOBAL-SCALE IONOSPHERE
by E.P.Szuszczewicz, R.A.Wolf, B.G.Fejer, R.W.Schunk and E.Roelof 38

Paper 39 withdrawn

NORTH-SOUTH TROPICAL F2 LAYER ASYMMETRY AT WEST AFRICAN AND PACIFIC LONGITUDES
by C.Davy, et al. 40

JOINT VHF COHERENT RADAR AND OBLIQUE SOUNDER OBSERVATIONS FOR SHORT TERM IONOSPHERIC FORECASTING
by P.S.Cannon, et al. 41

RADIO PROPAGATION CONDITIONS FORECAST ABOVE A LOW LATITUDE REGION
by H.Sizun and M.Lissillour 42

AN IONOSPHERIC MODEL FOR HF SKY-WAVE BACKSCATTER RADAR
by G.H.Millman, C.A.Bowser and R.W.Swanson 43

ACCURATE MEASUREMENTS OF THE TOTAL ANGLE OF ARRIVAL OF HF SKYWAVES
by Z.R.Jeffrey, P.T. Middleton and C.Winkler 44

IONOSPHERIC LIMITATIONS TO THE ACCURACY OF SSL ESTIMATES OF HF TRANSMITTER LOCATIONS
by L.F.McNamara 45

INFLUENCE DES VARIATIONS IONOSPHERIQUES SUR LES SYSTEMES HF A HAUTE FIABILITE UTILISANT DE GRANDES BASES
par C.Goutelard et J.P. Van Uffelen 46

* Abstract only

# TECHNICAL EVALUATION REPORT

BY

R.W. SCHUNK

Center for Atmospheric and Space Sciences
Utah State University
Logan, Utah 84322-4405

## 1.INTRODUCTION

The Electromagnetic Wave Propagation Panel sponsored a specialists' meeting on "Ionospheric Structure and Variability on a Global Scale and Interaction with the Atmosphere and Magnetosphere." The symposium was held in Munich, West Germany on 16–20 May 1988. The co-chairmen were L. Bossy (Belgium) and R.W. Schunk (U.S.A.), and the Program Committee consisted of Dr. H. Soicher (U.S.A.), Prof. T.B. Jones (U.K.), Dr. A. Brekke (Norway), Dr. G. Rostoker (Canada), Prof. R.D. Hunsucker (U.S.A.), Dr. K. Schlegel (Germany) and C. Senior (France). The meeting was divided into six sessions, including Solar-Terrestrial Relationships (session 1), Global Ionospheric Dynamics (session 2), Ionosphere-Magnetosphere Interactions (session 3), Ionosphere-Atmosphere Interactions (session 4), Ionospheric Irregularities (session 5), and Multistation/Multiparameter Observations and Systems Impact (session 6). In each session, there were both invited and contributed presentations. A total of 46 papers were presented during the five-day meeting. The meeting was ended with a round table discussion, where the session chairmen summarized their views on the current state of knowledge and the future direction of the field.

## 2. THEME

Hitherto, study of the aerospace propagation media tended to concentrate on localised phenomena or events. However, there is a range of phenomena which need for their elucidation observations and analysis on a global scale. Much of the local ionospheric behaviour is determined by coupling electrically to the distant ionosphere and solar wind, and coupling dynamically and electrodynamically to higher and lower levels of the atmosphere. Global multistation/multiparameter observations are often found to be necessary to provide the frame of reference for interpreting both stationary and propagating locally observed effects.

Communications, Navigation and Surveillance systems operating in/through the aerospace EM propagation environment are affected by the state/variability of the propagation media. An understanding of the complex global interaction would improve the means of predictability and assessment of localised phenomena and suggest methods for mitigation of adverse propagation conditions.

## 3. TECHNICAL EVALUATION

### 0.1. *SESSION I—SOLAR-TERRESTRIAL RELATIONSHIPS*

This session addressed the broad issues involving sun-earth coupling processes, including the effect of solar flares, the solar wind, and the interplanetary magnetic field on the ionosphere-atmosphere system. The session also dealt with more rigorous mathematical formulations of ionospheric dynamics and acoustic gravity waves. Four papers were presented in this session and the chairman was H. Rishbeth.

The first paper (1) was an invited review given by J. Lemaire, entitled "Solar-Terrestrial Relations: Flare and Solar Wind Effects." This paper emphasized that solar flares and solar activity in general affect the plasma and magnetic fields in the interplanetary medium. Both large-scale and small-scale inhomogeneities are produced in the solar wind, and the subsequent interaction of the solar wind with the earth's magnetic field acts to produce structure and variability in the ionosphere-atmosphere system. The discussion emphasized the need to include these effects in the modelling.

The following review paper by G. Rostoker (2) was entitled "Magnetospheric Substorms as a Signature of the Solar Terrestrial Interaction." This paper discussed the dynamics of substorms, and it was shown that changes in the interplanetary magnetic field (IMF) direction lead to substorms that last for a time ranging from a few minutes to a few hours. During changing IMF conditions, large amounts of solar wind energy penetrate into the magnetosphere. This energy is first stored as magnetic field energy in the magnetotail and as kinetic drift energy of particles in the plasma sheet and boundary layer. Ultimately, the energy is dissipated in the ionosphere-atmosphere system via field-aligned currents, energetic particle precipitation, and electric fields. There also was a discussion on the possibility of being able to predict the location and level of ionospheric perturbations during substorm. In this regard, it was noted that progress has been made in understanding substorm dynamics, but much work still remains before reliable predictions are possible.

Papers (3) and (4) dealt with mathematical models for ionospheric plasma transport and acoustic gravity waves, respectively. In the paper by K. Suchy, entitled "Balance Equations for Ionospheric Plasmas with Different Partial Temperatures", physically meaningful expressions for partial temperatures, partial energies, and partial heat fluxes were established and then the corresponding set of plasma transport equations were derived. These transport equations should provide a better description of the ionosphere for conditions when there are large temperature differences between the interacting species. In the paper by Miesen et al, entitled "Nonlinear Wave Equations for Low-Frequency Acoustic Gravity Waves," the characteristic nonlinear equations for low-frequency internal gravity waves were derived for an incompressible isothermal atmosphere because these waves affect the refraction index of the atmosphere, and hence, are relevant to electromagnetic wave propagation. The equations derived by the authors apply to the horizontal and time variations of the disturbance, and a discrete spectrum of group velocities for the waves was found.

### 0.2. *SESSION III--IONOSPHERE/MAGNETOSPHERE INTERACTIONS*

At the time of the meeting, after the abstracts were in hand, it made more sense scientifically to put session 3 ahead of session 2. This session dealt with all forms of ionosphere-magnetosphere coupling phenomena, including field-aligned currents, particle precipitation, electric fields, and ionospheric ion outflows. The session also covered the effects of magnetospheric processes on the ionosphere. Six papers were presented in the session, and it was chaired by C. Senior.

The first paper was a review given by M. Blanc (5), entitled "Ionospheric-Magnetospheric Coupling Processes." The review was primarily limited to a description of the theoretical work being done in this area, and it covered the high, middle and low latitude regimes. It was shown that magnetospheric-ionospheric coupling phenomena exist in all three ionospheric domains, but that they are strongest at high latitudes. It was also shown that there can be time delays between magnetospheric variations and ionospheric responses, which makes ionospheric forecasting difficult.

Dr. R.A. Hoffman (6) also presented a review on ionosphere-magnetosphere coupling, entitled "Observations of Ionosphere/Magnetosphere Interactions from the Dynamics Explorer Satellites." The observations presented were truly unique in that the data were obtained simultaneously from two satellites, one in the ionosphere and one in the magnetosphere. A striking finding was that during extremely quiet magnetic activity, the large-scale convection in the dark polar cap disappears, leaving only small scale structures. In general, it was found that the plasma convection pattern is a strong function of both the IMF and the substorm phase.

In her paper "Empirical Models of Electric Fields and Electrostatic Potential at High Latitudes from EISCAT Observations: Preliminary Results," C. Senior (7) presented average plasma convection patterns calculated from ion drift data obtained by the EISCAT incoherent scatter radar in Europe. The calculated patterns were consistent with the well-known two-cell structure of antisunward flow over the polar cap and sunward flow equatorward of the auroral zone. These patterns are needed in all ionosphere and thermosphere modelling studies of the high latitude domain.

Convection patterns were used in a case study by B. Watkins et al (8), entitled "Interplanetary Magnetic Field Effects on the Polar Ionosphere During Geomagnetic Disturbances." Specifi-

cally, the temporal variation of the polar $F$-region was modelled using a coupled magnetosphere-ionosphere system. The unique aspect of the modelling was related to the fact that the plasma convection and particle precipitation patterns varied systematically with the variation of the IMF, and the resulting ionospheric variations were shown to be significant.

The final two papers in this session dealt with wave excitation and plasma irregularities. K. Rinnert (9) in his paper entitled "Low Frequency Electrostatic Waves Observed in the Vicinity of an Auroral Arc" showed rocket data obtained in the vicinity of a faint auroral arc. The signals were ion cyclotron waves excited by ion beams on the equatorward edge of the electron precipitation region. In his paper entitled "Polar $F1$ Layer Lacuna Irregularities at Dumont d'Urville" P. Vila presented ionogram sequences to support his conclusion that the polar $F1$ layer lacuna irregularities are not always caused by the two-stream instability, but can be caused by plasma resonances driven by $F$ region parallel currents.

### 0.3. *SESSION II-GLOBAL IONOSPHERIC DYNAMICS*

This session covered all aspects of global ionospheric dynamics, including dynamical processes in the high, middle, and low latitude regimes. The dynamical features ranged from large-scale plasma convection at high latitudes to small-scale plasma bubble motion in the equatorial ionosphere. The session chairman was G. Prölss and ten papers were presented.

The first presentation was a review talk given by K. Rawer (11) entitled "Global Ionospheric Dynamics." The talk was tutorial in nature and covered the basic physics governing dynamical processes in the ionosphere. The talk emphasized the importance of the geomagnetic field in determining plasma drift features. It was also noted that neutral winds affect ionospheric motion at mid-latitudes, while electric fields are dominant at low and high latitudes.

The next presentation was also a review talk given by H. Rishbeth (12), entitled "Global Ionosphere Dynamics: A Review." Dr. Rishbeth reviewed various aspects of ionospheric structure and dynamics, including the effects of winds on the ionosphere, ionospheric storms, sporadic E, and both low and high latitude dynamical processes. The emphasis was on our ability to model the ionosphere and both empirical and physical models were discussed.

The paper by D. Anderson and J.M. Forbes (13) on "A Fully Analytic, Low and Mid-Latitude Ionospheric Model" was also concerned with ionospheric modelling. In particular, a new analytical model of the ionosphere was described. This model generates electron density profiles for a range of solar cycle conditions (minimum, moderate and maximum) and for different seasonal conditions (winter, summer and equinox). Differences between the new model and previous analytical models were discussed.

K. Schlegel et al (14) discussed the formation of plasma bubbles in their paper entitled "Global Morphology of Plasma Bubbles in the Low Latitude Ionosphere." Measurements of the electron density and temperature in and around plasma bubbles were presented. The bubbles were observed near the magnetic equator, and it was found that the temperature in the bubbles can be both higher and lower than the surrounding plasmas, which was a surprising result.

A paper entitled "Simulation of the Polar Cap $F$ Region Ionization Using an Experimental Convection Electric Field" was presented by C. Taieb (15). This paper discussed the results obtained from a numerical model of the polar ionosphere. Measured electric fields were used to obtain a plasma convection pattern, and the ionospheric response to this pattern was calculated and compared to electron densities measured by the EISCAT incoherent scatter radar. There was excellent agreement between measured and calculated ionospheric parameters.

G. Wrenn et al (16) presented "Average Diurnal and Seasonal Profiles of Ionospheric Storms." These authors noted that the forecasting of Maximum Usable Frequencies for short-wave radio communication is not reliable because of $f_oF_2$ reductions during ionospheric storms. An extensive ionosonde data base was assembled from five mid-latitude sites over a complete solar cycle with the hope that the data might shed light on the cause of ionospheric storms. The characteristics of the data were presented and modellers were encouraged to study the data.

"Medium Scale Structure of the $F$-Region" by A. Paul (17) was a presentation that emphasized the need to make more frequent ionosonde measurements. Currently, routine ionosonde measurements are made on an hourly basis. However, digital ionograms were presented supporting the hypothesis that acoustic gravity waves cause important short term $F$-region variations. The gravity waves appear to be present at all times, and therefore, more frequent ionospheric measurements are needed.

Paper number (18) by P. Kossey and J.E. Rasmussen was entitled "VLF/LF Radiowave Studies of the Structure of the Equatorial, Mid-Latitude and Polar Ionosphere Below 100 km." Data obtained in the polar cap during solar proton events were presented, and the effect of such ionospheric disturbances on VLF/LF/HF/VHF military systems was discussed.

Y. Tulunay (19) presented a paper entitled "The Middle and High Latitude Ionosphere at About 550 km Altitude." Ariel 4 satellite data were presented in order to describe the general morphology of the electron density in the mid-latitude trough, the polar cap depletions, and the auroral enhancements. At equinox and for quiet magnetic activity, the maximum electron densities observed in the northern and southern hemispheres were found to differ significantly.

The final paper of the session was by R.W. Schunk and J.J. Sojka (20), entitled "Modelling Ionospheric Density Structures." From the model study it was found that *F*-region plasma density structures in the polar region can be created by a variety of processes, including structured electric fields, structured precipitation in the auroral oval or in sun-aligned polar cap arcs, plasma instabilities, and temporal variations of the IMF. It was also found that the lifetime of large-scale ionospheric density structures depends on the seasonal and solar cycle conditions.

### 0.4. *SESSION IV—IONOSPHERE/ATMOSPHERE INTERACTIONS*

This session covered a broad range of ionosphere-atmosphere coupling phenomena, including large-scale momentum and energy coupling at high latitudes, the effect of non-thermal plasmas on the coupling processes, the influence of TID's and gravity waves on the coupling, and the influence of ionospheric conductivity. There were nine papers presented and the session chairman was G. Rostoker.

The first paper was a review of the recent ionosphere-atmosphere model studies by D. Rees and T.J. Fuller-Rowell (21), entitled "Seasonal and Universal Time Variations of the Geomagnetic Response of the Thermosphere and Ionosphere." Using a fully-coupled self-consistent model of the ionosphere-thermosphere system, the authors presented a series of simulations for moderate and disturbed geomagnetic activity levels, high and low solar activity, and for various seasonal conditions. The complicated coupling processes were pointed out as was the great variability of the ionosphere-atmosphere system at high latitudes. The results emphasized the difficulties associated with developing computationally fast empirical models of the high latitude region.

G. Prölss (22) also presented a review talk on "Ionosphere-Atmosphere Interactions," but this presentation emphasized experimental results rather than theory or modelling. An excellent discussion was given on the inability of existing models to explain certain observed thermospheric effects, such as ionospheric storms.

K.J. Winser et al (23) presented a paper on "Ionosphere-Thermosphere Coupling at High Latitudes: Aspect Angle Dependence of Non-Thermal Plasmas." EISCAT radar measurements were presented that tend to confirm theoretical predictions of non-Maxwellian ion velocity distributions in ionospheric regions where the plasma convection speed exceeds 1 km/s. The non-thermal features were shown to affect ionosphere-atmosphere momentum and energy coupling.

In a paper by A. Brekke and C. Hall (24), entitled "Quiet Time Conductivities of the Auroral Ionosphere," EISCAT radar data were used to derive auroral conductances and the results were compared to previous work. It was noted that different authors have obtained different results, and some of the discrepancies can be traced to the use of different collision frequency models.

D. Figueroa and H. Kohl (25) discussed the "Calculation of the Neutral Gas Temperature and Velocity Vector Using EISCAT CP-3 Data." Using radar data as an input, the energy equation for the ions and the momentum equation for the neutrals were solved in order to obtain the temperature and horizontal velocity of the neutral gas. The calculated neutral gas velocities showed the strong influence of plasma convection at high latitudes.

Paper (26) was by I.W. McCrea et al, entitled "Ion Heating Events Observed by the EISCAT Radar." EISCAT radar data were examined in an effort to identify times when there were elevated ion temperatures. The majority of the enhanced ion temperatures were observed in the evening and early morning sectors. The existence of enhanced ion temperatures at these times was attributed to the occurrence of lower plasma densities, and hence, decreased ion-neutral coupling, which results in larger ion-neutral drift velocity differences. The latter, in turn, leads to increased frictional heating rates.

Kamp et al (27) presented a paper entitled "On the Influence of the Earth's Magnetic Field on Ionospheric Acoustic Gravity Waves." The authors derived a generalized dispersion relation

for all modes, including acoustic gravity waves and electromagnetic waves. They found that the ionospheric conductivity results in a strong temporal damping of the acoustic gravity waves.

Gravity waves were also discussed by H.-P. Mauelshagen and K. Schlegel (28) in a paper "Atmospheric Gravity Waves in the Auroral Source Region of TID's Studied with the EISCAT CP2-Data." These authors studied the source of gravity waves using high-resolution EISCAT radar data. The different propagation directions of the waves were then examined with respect to their position in the auroral oval and with respect to substorm phases, which allowed for an identification of the source region.

The final paper in the session by U. Inan was withdrawn, and it was replaced with a paper given by M. Cutolo (29), entitled "Artifical Ionization in Polar Aurora Attained in the Ionosphere by Radio Waves Emitted from the Ground."

### 0.5. *SESSION V—IONOSPHERIC IRREGULARITIES*

This session was concerned with small-scale ionospheric structure and plasma irregularities. The papers presented covered both man-made and natural plasma irregularities. The discussions concerning natural irregularities included those occurring in the polar cap, auroral $E$ region, and low latitude ionosphere. The effect of irregularities on electromagnetic wave propagation was also discussed. The session chairman was A. Brekke and seven papers were presented in this session.

Professor J. Fejer (30), who is one of the world's leading experts on plasma irregularities, presented a comprehensive review on "Ionospheric Irregularities Due to Powerful HF Radio Transmissions." Professor Fejer discussed all aspects of the phenomena that play a role in the production of irregularities by powerful HF transmitters. He discussed thermal-focusing of radio waves and the production of irregularities by thermal parametric instabilities as well as parametric instabilities in which the pondermotive force dominates over thermal forces. Professor Fejer also discussed the practical applications associated with artificially-induced ionospheric irregularities.

"A Review of Radio Studies of Auroral $E$-Region Ionospheric Irregularities" was presented by C. Haldoupis (31). The discussion centered on recent radar studies of auroral irregularities at different frequency bands in the VHF and UHF range. It was noted that the original work on 'auroral' $E$-region irregularities was heavily influenced by the work done on 'equatorial' irregularities because in both regions there are strong horizontal currents carried by $\mathbf{E} \times \mathbf{B}$ drifting electrons. However, recent auroral observations have shown that there are several other physical processes operative in the auroral region that can produce $E$-region irregularities. Professor Haldoupis also summarized the major unresolved problems in this field.

Professor K.C. Yeh (32) presented an excellent paper on "Ionospheric Irregularities and Their Effects on Electromagnetic Waves Propagating Through Them". He noted that the ionospheric medium is dispersive and in the presence of irregularities additional wave scattering can take place. Therefore, a radio wave propagating through an environment containing irregularities can experience fading and scintillations. He also noted that it is now well-known irregularities strongly depend on the phase of the solar cycle, the geomagnetic latitude and local time.

Paper (33) by Cerisier et al was on "Electrostatic Turbulence in the High Latitude Ionosphere." Electric field data from the low altitude polar orbiting satellite Aureol-3 were analyzed to study the characteristics of electrostatic turbulence. The turbulence features were then used to determine the possible sources of free energy needed to excite the turbulence, including field-aligned currents, particle precipitation and electron density gradients.

The paper by Aarons et al (34) was on "Auroral and Sub-Auroral $F$ Layer Irregularity Studies in the Northern and Southern Hemispheres During the Equinox Transition Study." Data on $F$-layer irregularities obtained during the 1984 Energy Budget Campaign were presented. It was found that for sub-auroral latitudes there was a dramatic difference between the irregularity intensity during a period of magnetic quiet at the beginning of the campaign and a period of magnetic quiet in the recovery period of a magnetic storm. The authors suggested that the ring current played a significant role in this time period.

Hanuise et al (35) presented a paper on "Small-Scale Irregularities and Global Plasma Dynamics in the High Latitude Ionosphere." The authors described a measurement program, based on new sophisticated HF coherent radars, called Polar Region Ionosphere Structure Monitor (PRISM). With the new system one can derive a two-dimensional map of plasma convection and electric fields over a region as large as 2 million square kilometers, with a good spatial resolution. It therefore offers a unique opportunity to study the formation and dynamics of small-scale ionospheric structures.

The final paper in this session was by K. Bibl (36), entitled "Examples of Meteorological Behavior of the Ionosphere". One example shown was called the 'European Anomaly' by the author. In fixed frequency ionosonde recordings, this feature was discovered as direct oblique echoes from a hole in the ionosphere over the Alps. The author presented maps of this meteorological ionospheric phenomenon in terms of contours of plasma density over Europe obtained from ten stations. The reason for this ionospheric feature is unknown.

### 0.6. *SESSION VI—MULTISTATION-MULTIPARAMETER OBSERVATIONS AND SYSTEMS IMPACT*

The final session involved a discussion of the efforts underway to determine the global-scale characteristics of the ionosphere using data obtained simultaneously from multiple ground-based sites and from satellites. The emphasis here was on ionospheric forecasting. This session was also concerned with the effect of the ionosphere on military systems. There were ten papers presented in this session and the chairman was R.W. Schunk.

B. Reinisch (37) described a new generation of ionosondes that is currently being employed in his paper entitled "Multistation/Multiparameter Observations with a Network of Digital Ionosondes." By the end of this year, forty digital ionosondes spread around the high latitude region will provide a simultaneous data set of ionospheric parameters ($f_oF2$, $f_oF1$, $f_oE$, etc.) that are automatically scaled in real time. All the data can be remotely accessed by telephone links. These data should be invaluable for obtaining the 'instantaneous' plasma convection pattern and for studying gravity wave propagation.

A paper entitled "Joint SABRE and Oblique Sounder Observations for Short Term Ionospheric Forecasting" by Cannon et al (38) was the second paper in this session. In this presentation, results from a series of experiments employing auroral radars and oblique ionospheric sounders were described. The idea of the experiments is to obtain the location of the polar cap and auroral oval boundaries in real time so that this information can be used in physical models of the ionosphere. This would greatly aid ionospheric prediction at high latitudes.

Dr.E. Szuszczewicz (39) presented a paper on "Ionospheric Measurements from a Worldwide Chain of Stations." In particular, he described the SUNDIAL program, which is a coordinated observation program involving 70 ground-based stations coupled with an extensive modelling effort. The overall goal of the SUNDIAL program is to study the flow of energy from the sun to the earth's upper atmosphere in an effort to develop ionospheric prediction models.

Paper (40) was by D. Anderson and H. Carlson, entitled "The Effect of the Ionosphere on Air Force Systems."

A paper entitled "North-South Tropical $F2$ Layer Asymmetry at West African and Pacific Longitudes" was presented by P. Vila et al (41). In this presentation, data from the 1986 equinox SUNDIAL campaign were discussed with regard to Appleton Anomaly features. The intertropical fountain and Appleton density peaks were shown and discussed in relation to the effects of $\mathbf{E} \times \mathbf{B}$ drifts and neutral wind drag.

H. Sizun et al (42) talked about "Radio Propagation Conditions Forecast Above a Low Latitude Region" and G.H. Millman et al (43) presented a paper entitled "An Ionospheric Model for HF Sky-Wave Backscatter Radar." In the latter paper, it was noted that the geographic coordinates of a target detected by a line-of-sight radar can be derived from the measurement of the elevation and azimuthal angles and the time delay (radar range). However, the determination of the target location requires that the virtual height of reflection of the radar transmissions also be known. An ionospheric model in conjunction with a locally recorded ionogram can be employed to estimate the virtual height. The authors then described a relatively simple ionospheric model applicable to the HF backscatter radar over-the-horizon identification problem.

The paper by Jeffrey et al (44) on "Accurate Measurements of the Total Angle of Arrival of HF Skywaves' and the one by McNamara (45) on "Ionospheric Limitations to the Accuracy of SSL Estimates of HF Transmitter Locations" were combined into a single oral presentation at the time of the meeting. The presentation was given by L. McNamara. He discussed the Single Station Location (SSL) approach to the problem of locating HF transmitters. He noted that the accuracy and resolution of good SSL systems are now such that the major errors in position estimation arise from an inadequate knowledge of the ionosphere along the circuit. Although ionospheric models are currently being used, further theoretical work is needed to improve the present empirical profile

shapes and to introduce more realistic horizontal gradients into the models. He also stressed the need to identify sporadic E along the circuit.

The final paper in the session was by Goutelard et al (46), entitled "Influence des Variations Ionospheriques sur les Systems HF a Haute Fiabilite Utilisant de Grandes Bases."

# Solar-terrestrial relations : flare and solar wind effects

J. Lemaire
Institute for Space Aeronomy, 3 ave. Circulaire, B-1180 Bruxelles, Belgium

## Abstract

The existence of Solar Terrestrial Relations is clearly shown during solar flare events. During these catastrophic events the enhanced flux of XUV solar radiation as well as the sudden outburst of energetic solar cosmic ray particles induce a series of well identified effects in the Earth's magnetic field, in the terrestrial ionosphere and in the upper atmosphere. These geophysical effects and their solar origin will be reviewed in the first part of this article.

Solar activity also perturbs the distribution of plasma and magnetic fields in the corona and in the interplanetary medium. It produces large-scale (low frequency) and small-scale (high frequency) inhomogeneities (perturbations) are produced in the expanding solar wind flow which then interacts with the Earth's magnetic field. The resulting variability of the impinging solar wind plasma triggers additional geophysical effects in the magnetosphere, in the terrestrial ionosphere and in the upper atmosphere. Some of the geophysical consequences of this non-stationary interaction of the solar wind with the geomagnetic field, and, of impulsive penetration of small scale solar wind plasma irregularities into the magnetosphere will be reviewed in the second part of this article (e.g. polar cusp fields, currents, and related ionospheric irregularities).

## 1. Introduction

First evidence that the Sun is not a changeless, immutable light and heat source, came very early from the observations of changing spots on the Sun's surface. These sunspots appear to develop at unforeseable places, and they disappear after a lifetime varying between two and a hundred days. The number of sunspots is variable over a period roughly equal to two times eleven years : the period of the solar activity cycle. (see fig. 1).

That the Sun is changing and can become suddenly very active over periods of minutes, was discovered in 1859 by R.C. Carrington and R. Hodgson when they observed for the first time what is now known as a "solar flare" ; they described what they observed in the midst of a large sunspot group, as "two patches of intensely bright and white light which suddenly break out".

## 2. Solar Flares

Flares are the most dramatic eruptions which release over their short lifetime (5 to 100 minutes) a total energy ranging between $10^{29}$ ergs (in the case of subflares) and $10^{32}$ ergs (in the case of a Flare of class 4. Even relatively frequent small flares expend an amount of energy equivalent to millions of hydrogen bombs.

The rate of occurence per day (E) of solar eruptions varies in phase with the relative sunspot number (R ; also called Wolf number) :

$$E = 0.04\ R(d^{-1})$$ (Ref.1).

One of the most studied superflares is that of November 12, 1960. Most recently a series of six large flares was observed during the period of 5-7 February 1986. This series of events triggered a very intense geomagnetic storm and caused disruptions in electric power lines (Refs. 3, 4).

### A) Origin of flares and electromagnetic radiation

The origin of solar flares remains an unresolved physical problem ; however, since magnetic field intensities are always high ($10^{-2}$ to 1 Tesla) in solar flare areas, magnetohydrodynamic geophysicists speculate that these explosions result from the conversion of magnetic energy into kinetic energy. This should, in principle, occur along neutral lines or at neutral points where the magnetic field intensity reverses direction : magnetic field lines are assumed to reconnect (or to merge) at these places ; this is supposed to give explosive flare events. Most of these Reconnection models, however, are steady state or quasi-steady ; induced (A.C.) electric fields are ignored in "static" type models for these highly (non-stationary) explosive events! Therefore one can but wonder what is the actual relevance of current reconnection models in the case of highly non-stationary hydrodynamical solar flare events. But other physical mechanisms should not be excluded to explain the origin of solar flares.

It is known that the explosion of energy takes place in the chromosphere and the low altitude region of the corona where the temperature increases from the minimum photospheric value (about 4,000 K) to more than a million degrees in the solar corona over an altitude range of less than $0.01\ R_S$ ($R_S$ is the solar radius : $6.9 \times 10^8$ m).

The gigantic energy release is exported by a brilliant burst of visible light (H-alpha) and of electromagnetic waves ranging from X-rays to radio waves, by ions and electrons accelerated to more than half the speed of light, and, by clouds of ionized gas that sweep through interplanetary space at 700-800 km/s. Recently, the study of high-energy flares has been greatly advanced by spacecraft observation of gamma-ray spectra and neutron fluxes (Ref. 25).

The solar X-ray, like the solar radio emissions, originate from both thermal and non-thermal processes that take place primarily in the solar corona. Non-thermal X-ray emissions occur mainly at short wavelengths less than about 0.5 nm corresponding to photon energies greater than 2 keV (Ref. 7, 36). Gamma ray line observations of fluxes from SMM and Hinotori satellites do not support the two phase concept of solar flare development.

**Geophysical effects of electromagnetic radiation of solar flares**

Among the most spectacular geophysical effects produced by bursts of hard X-rays are Sudden Ionospheric Disturbances (SID). These ionospheric disturbances are observed in the sunlit hemisphere and arise simultaneously with visual flare observation H-alpha. The geophysical effects result from the enhanced ionization of the E- and D-regions of the terrestrial ionosphere (Ref. 36). SIDs and the subsequent recovery of the ionosphere have a time duration somewhat longer than the flare duration ; generally from a few minutes to an hour, with more rapid rise than decay.

Several SID observations are routinely identified and digitized. Let us just mention a few of the most important of them

1) Solar Flare Effects (SFE), or Crochets, are observed as a small hook on magnetometer records. They are caused by the magnetic field response to increased current flow in the E-region due to electron density enhancement induced by flare X-rays.

2) Short Wave Fadeouts (SWF), are observed from signal strength records of shortwave receivers (3 to 30 MHz). Signals from sweep-frequency ionosondes may be completely absorbed as a consequence of the enhancement of electron density induced by solar flare X-rays in the D-region.

3) When an SID occurs, the vertical height of reflection is suddenly lowered, and the path length shortens. Ionospheric disturbances resulting in this changed path length are known as Sudden Phase Anomalies (SPA).

4) Sudden Enhancement of Atmospherics (SEA), are observed as an increase in signal strength on wideband equipment operated to detect electromagnetic emissions from lightning at frequencies ranging between 10 and 50 kHz

There are also other types of SID's and geophysical effects produced by the bursts of electromagnetic radiation emitted by the Sun at the onset of flares. A comprehensive list of these effects is given for example in the Handbook of Correlative Data (Ref.2).

C) **Corpuscular radiations emitted by flares and their geophysical effects**

1) Solar Energetic Particles (SEP)

Significant fluxes of electrons, protons and multiply charged ions are accelerated at the Sun during energetic solar flares. Some of these MeV particles are subsequently detected within the interplanetary medium and in the Earth magnetosphere.

The largest flares release protons and heavy ions with energies up to 10 GeV i.e. overlapping in energy with the galactic cosmic rays (see fig. 2). The heavy metallic ions with low first ionization potentials have been found to be relatively more abundant in the streams of Solar Energetic Particles, compared to their abundances in the photosphere (Ref. 3).

The penetration of these energetic particles in the Earth's atmosphere is easiest at higher latitudes. Meeting weaker resistance from the geomagnetic field over the polar caps these protons crash into the atmosphere, exploding the molecules of Oxygen and Nitrogen that they hit into many energetic shower particles which reach the ground over a wide area.

The penetration of Solar Energetic Particles (SEP) into the polar cap ionosphere enhances considerably the ionization in the D-region. Consequently Solar Proton Events (SPE) are marked by strong radio absorption which is observed 15 to 100 minutes after the visual flare brightening. These absorption eventswhich are called Polar Cap Absorptions (PCA), may persist for 1 to 6 days.

By timing the arrival of particles of different speeds, satellite measurements have shown that while all the particles have been produced simultaneously in the flare and have travelled the same distance, the distance covered is generally much larger than that from the Sun to the Earth ; this suggests a roundabout trajectory resulting from numerous "collisions" of the particles with irregularities in the spiral interplanetary field (pitch angle scattering).

Both the intensity and spectrum of solar cosmic protons and heavy ions depend on the relative position of the Earth and the flare on the Sun. The actual amount of particles bombarding the magnetosphere and the atmosphere of the Earth depends therefore on the interplanetary conditions at the time of the flare. These conditions are, however, highly variable and unpredictable. This effect may lead to a variation by factors as large as 100 in the observed flux, at different points around the orbit of Earth from the same flare (Ref. 6).

2) Interplanetary shock waves

Solar energetic particles can reach the Magnetosphere of the Earth in less than one hour, but they do not contribute much to the thermal plasma density forming the bulk of the solar corona and solar wind. Within one or two days after a burst of SEP, particles from the solar plasma begin to arrive at the edge of the geo-magnetosphere i.e. at the magnetopause. The solar wind bulk velocity is then enhanced from less than 400 km/s (prevaling during quiet solar wind conditions and in interstream regions) up to 700-800 km/sec in the post- shock plasma ejected out of the corona after large solar flares.

When such a shock front hits the magnetopause the dayside region of the magnetosphere is "compressed". The sudden impulse (s.i.) or the storm sudden commencement (ssc) observed in magnetograms are the manifestation at ground level of this magnetospheric compression due to the enhanced solar wind pressure (see fig.3).

When the interplanetary magnetic field is directed southward at the time of this interaction the magnetic perturbation is likely to develop into a geomagnetic storm with a main phase and a recovery phase lasting for several hours and sometimes longer than a day.

During the storm sudden commencement phase, the mid and low latitude values of H (the horizontal component of the geomagnetic field) are simultaneously increased worldwide, typically by a few tens of nanoteslas (gammas). This state of enhanced H may persist for a few tens of minutes during the initial phase. During the following main phase of the magnetic storm, H is depressed below its prestorm value typically by several tens to a few hundred nanoteslas for many hours. The recovery phase, or return to prestorm state, usually requires a few days. These two latter phases are due to

the great enhancement and gradual decay of a large scale diamagnetic ring current set up within the magnetosphere.

Because of their relatively low energy (~ keV) the solar wind protons enter preferentially into the magnetosphere and down into the earth's atmosphere along high latitude magnetic field lines. The easiest and most direct access for solar wind particles is via the polar cusps which are formed by all magnetic field lines parallel to the magnetopause or traversing it (see fig.4) (Refs.4,5).

The whole polar caps are also invaded by the thermal solar wind proton forming there what has been called the Polar Rain (Ref.34). The polarity of the interplanetary magnetic field controls very effectively the access to the Northern and Southern polar caps (Ref.37).

## 3. Solar wind and its interaction with the Earth

Besides geophysical effects triggered from time to time by emissions of solar flare radiation and corpuscules, as recalled in the previous section, there is a whole class of solar terrestrial relationships linked to the existence of the solar wind interacting with the Earth's magnetic field, even in absence of flare events. Indeed, geomagnetic storms are often triggered by collisionless shocks which are not related to solar flare events, but which are caused by disappearing filaments, coronal holes (Refs. 9,29) or other unknown or yet unclear causes. Furthermore, as a consequence of their unfavorable location on the solar disk, some flares do not produce significant effects at the Earth.

### A. Origin and models of the solar wind

The existence and origin of the solar wind can be explained from different alternative points of view. In Appendix 1 we have followed the approach of kinetic theory which shows more explicitely and directly than any fluid approach why protons are accelerated out of the solar gravitational potential and why they obtain supersonic velocities at large heliocentric distances.

It is now well established that in the solar wind, there are two flow regimes : (i) high speed streams separated by regions where the bulk velocity is smaller than 400 Km/s, and, where the density is comparatively large ; (ii) the interstream slow solar wind which is mainly observed at boundaries of magnetic sectors and in vicinity of magnetic neutral sheets.

High speed streams are observed on both sides of interplanetary current sheets and IMF sector boundaries. They originate in regions of the solar corona where the plasma density and temperature are significantly depressed, and, where the magnetic field lines are "open", i.e. extend from the Sun far out beyond the orbit of the Earth ; these low temperature regions appear darker in X-Ray images of the corona and are called coronal holes.

Among the many surprising features observed in fast streams is the unexpected presence of narrow field aligned beams of suprathermal electrons which have been called "strahl electrons" (Ref. 19). These "strahl electrons" form a peculiar subpopulation besides the "core" and "halo" electrons which are observed in slow solar wind regimes (Ref.39). The origin of "strahl electrons" is still unknown !

The heavy ion abundances and temperatures also differ significantly in the two flow solar wind regimes (Ref.40). Many of these peculiarities remain unexplained and cannot be accounted for by current hydrodynamical models unless some ad hoc assumption is invoked to make them work and fit the observations. Ad hoc intensities of MHD waves and their expected dissipation in the Chromosphere and Corona up to 10 $R_S$, is the 'Deus ex machina' usually invoked !

### B. Interaction of the solar wind and magnetosphere

Most large-scale solar wind models are based on simplifying assumptions such as : (i) stationary flow and (ii) the absence of small scale plasma irregularities i.e. near uniformity of the plasma. Dozens of models for slow solar wind and high speed streams are based on these two easy and convenient assumptions. For more than two decades this has lead other theoreticians to consider that the interaction between the solar wind and the geomagnetic field could satisfactorily be described as a stationary or quasi-steady-state interaction. Tens of magnetospheric models produced since 1960 are indeed based on these easy and convenient assumptions i.e. (i) stationary flow and (ii) uniformity of the flow impinging on the magnetopause surface.

The deflection of a stationary laminar solar wind flow around the terrestrial magnetic obstacle can to some extent be simulated in the Laboratory ; when a neutral plasma stream, drifting across magnetic field lines, with a velocity $\underline{E} \times \underline{B}/B^2$, penetrates into the region of higher magnetic field intensity B, the component of its bulk velocity parallel to the gradient of B is reduced adiabatically. Indeed as a result of conservation of magnetic moment, the translation kinetic energy of all drifting charged particles is adiabatically converted into gyromotion. This leads to specular reflection of the particles against the fictive surface of the magnetosphere. The adiabatic deceleration of plasma streams in regions with increasing magnetic field intensity has been first proposed by Chapman and Ferraro (Ref.44) and demonstrated experimentally in the laboratory by Demidenko et al. (Ref.45) (see also Ref.23 and references therein).

Closed and open steady state models of the magnetosphere, like those schematically illustrated in fig.5a, have had their supporters for a long time. None of these stationary interaction models has, however, been able to account for all observations collected from space as well as from modern ground-based geophysical stations.

It was in 1976, at a meeting in Amsterdam, that the idea of non- stationary and patchy interaction of solar wind with the magnetosphere was emphasized, probably for the first time (see Refs. 20 and 21).

The interaction model put forward in Ref.20, assumed that small scale eddies are formed by MHD instabilities at "funnel-shaped identations" along the magnetopause surface. According to the other non-steady state interaction model proposed in Ref.21 the observed small scale plasma irregularities are not necessarily produced locally in the magnetopause region : it was argued that there are already plenty of such plasma irregularities in the non-uniform solar wind rushing into the

geomagnetic field ; the densest of these plasma elements or plasmoids are able to penetrate deeper into the magnetosphere than those with lower momentum flux density (Refs. 21, 22, 23). This second scenario, illustrated in fig.5b, is now known as the Impulsive Penetration (IP) theory of solar wind plasma density into the magnetosphere.

From high resolution interplanetary magnetic field measurements it was clear already in the 60's that the solar wind is non-stationary and non-uniform most of the time over periods as short as 5 seconds or distances of 2000 km (i.e. a few mean proton gyroradii). High resolution plasma measurements in the solar wind now available have confirmed this expectation inferred in 1976 from high resolution IMF observations.

Since the solar wind plasma is patchy, its momentum flux density over the surface of the magnetopause is neither uniform, nor stationary. Time dependent and non-uniform boundary conditions imposed from the outside by the variable and patchy solar wind, force the magnetospheric plasma, electric and magnetic fields to be non-stationary over periods of time of seconds.

This implies that the interaction between the solar wind and the geomagnetic field can not be described in the framework of D.C. electromagnetic field theory (i.e. the magnetostatic and electrostatic theory). The rapidly changing solar wind imposes a non-stationary description for this interaction : in other words A.C. electromagnetic effects have to be taken into account as claimed in Refs. 42, 43. This conclusion is similar to that reached in § 2A concerning the relevance of stationary reconnection models used to describe highly non-stationary explosive flare events.

The Impulsive Penetration Model briefly described in Refs. 21, 22 and 23, belongs to this second category. When diamagnetic solar wind plasma irregularities "rain" into the magnetosphere, like droplets into a water pound, they perturb not only the local plasma density, the ambiant temperature and ion composition, but, they also change continuously the local magnetic field intensity and its direction ; they perturb the local electric field, as a consequence of their relative motion with respect to the background magnetospheric plasma.

If A.C. electromagnetic effects are essential in the study of solar wind-magnetosphere interaction models, they cannot be described in the framework of slowly varying D.C. magnetostatic field line distributions like those generally used to sketch the topology of the magnetosphere or of Flux Transfer Events.

### C. Geophysical effects resulting from impulsive penetration of plasmoids in the geomagnetic field

Intrusion of the most impulsive diamagnetic plasma blobs is controlled by the orientation of the interplanetary magnetic field. This is a consequence of the dipole-dipole interaction between the magnetic moment of the entering plasmoids and the magnetic moment of the Earth (Refs. 22, 23, 41).

Furthermore, while they are convected across magnetic field lines in the magnetosheath and in the magnetosphere with an injection drift velocity, $\underline{E} \times \underline{B}/B^2 = V$, the particles forming the penetrating plasmoids are eventually slowed down ; their injection (translational) energy, $1/2\ mV_\perp^2$, is converted into enhanced gyro-motion because of conservation of their magnetic moment, $\mu = \frac{1}{2} mv_\perp^2/2B$. Their penetration velocity, $\underline{V}$, as well as $\underline{E}$, the convection electric field inside the plasmoids, eventually vanish when they have reached a place where the local magnetic field intensity B is such that all their translational energy has been transferred adiabatically into thermal motion perpendicular to the magnetic field direction (Ref. 23).

In addition to this adiabatic braking or slowing down of all particles forming the invading plasmoids, non-adiabatic (irreversible) dissipation of the injection energy takes place via Ohmic (Joule) heating in the dayside Cusp or Cleft Ionosphere (Ref. 21).

This distant energy dissipation heats the plasma at the low altitude feet of all the magnetospheric field lines which are crossed by penetrating plasma irregularities. This ionospheric heating of the plasma in the dayside clefts has been observed and is discussed in Refs. (30, 35). This is one of the clearly identified geophysical effects resulting from the solar wind-magnetosphere interaction.

Another geophysical consequence of Impulsive Penetration theory is the transfer of momentum to the ionospheric plasma in the high latitude trough regions and also to the magnetospheric plasma flowing around intruding plasmoids. This is rather analogous to the flow of any fluid along the surface of a penetrating object (Ref. 22, 23). The resulting poleward ionospheric motions have been identified and discussed in Refs. 31, 32 and 33.

The field-aligned propagation and dispersion of the ions and electrons forming the intruding plasmoids have also been observed and reported in Ref. 5. The description given for these magnetosheath particle injection observed in the dayside cleft fits nicely in the Impulsive Penetration scenario. They form an additional class of geophysical phenomena which are consequences of the non-stationary interaction between the solar wind and the Earth's magnetosphere illustrated in fig. 5b.

Other geophysical effects like field-aligned current systems and non-steady state convection flow patterns observed in the polar caps can be interpreted in terms of the theory of Impulsive Penetration of solar wind plasma irregularities in the magnetospheric tail lobes, when the Interplanetary Magnetic Field has a Northward component. These NBZ effects are described in Ref. 41.

There are many more geophysical effects which are linked directly or indirectly to solar wind conditions and to their time variations (see for instance Ref.24). For up to date and comprehensive reviews, see for instance Refs. 26, 27 and 28. All these phenomena are part of the wide ensemble of solar-terrestrial relationships. Some of these additional solar-terrestrial relationships are discussed in the next article and in other papers in these proceedings.

## 4. Conclusions

Let us conclude this necessarily incomplete review of solar-terrestrial relationships by emphasizing how much the Sun-Earth system is a floppy coupled physical system. The links are basically electromagnetic and corpuscular radiation at all wavelengths and at all energies. The intensity of these different fields of radiation are not only variable over periods comparable to the 11 years

solar cycle, but they also change drastically over time scales of minutes and even seconds.

A lesson of this scientific endeavour could then be : although the assumptions of stationary and uniform flow have always been tempting and convenient for traditional modelers, these easy assumptions happen rarely to be applicable to the ever changing Sun, Solar Wind, Magnetosphere, Ionosphere, Upper Atmosphere ... and, of course, *Solar Flares, Flux Transfer Events, or Magnetic* Substorms.

## 5. References


1. TEN BRUGGENCATE, P., Veröffentl. Göttingen Nr.104, 1953.
2. KING, J.H. (Ed.), Handbook of correlative Data, NSSDC-71-05, GSFC, Greenbelt, MD, 1971.
3. MEYER, J.-P., The baseline composition of solar energetic particles, Ap. J. Suppl. Ser., 57, 1985, 151-171.
4. PETERSON, W.K., Ion injection and acceleration in the polar cusp, in "The Polar Cusp", eds. J. Holtet and A. Egeland, D. Reidel Publ. Co., Dordrecht, 67-84, 1985.
5. CARLSON, C.W. and TORBERT, R.B., Solar wind injection in the morning auroral oval, J. Geophys. Res., 85, 1980, 2903-2908.
6. SIMNETT, G.M., Solar cosmic radiation during August 1972, *Space Sci. Rev.*, 19, 1976, 579-610.
7. McGUIRE, R.E., The composition, propagation and acceleration of energetic solar particles, Rev. Geophys. Space Phys., 21, 1983, 305.
8. ALLEN, J.H., Major magnetic storm effects noted, EOS, 67, 1986, 537.
9. GARCIA, H.A. and DRYER, M., The solar flares of February 1986 and the ensuring intense geomagnetic storm, Solar Phys., 109, 1987, 1987.
10. *PARKER, E.N., Dynamics of the interplanetary gas and magnetic fields*, Astrophys. J., 128, 1958, 664-676.
11. CHAMBERLAIN, J.W., Interplanetary gas. II. Expansion of a model solar corona, Astrophys. J., 131, 1960, 47-56.
12. LEMAIRE, J. and SCHERER, M., Kinetic models of the solar wind, J. Geophys. Res., 76, 7479-7490.
13. LEMAIRE, J. and SCHERER, M., Kinetic models of the solar and polar winds, Rev. Geophys. Space Phys., 11, 1973, 427-468.
14. HUNDHAUSEN, A.J., BAME, S.J., ASBRIDGE, J.R. and SYDORIAK, S.J., Solar wind proton properties : Vela 3 observations from July 1965 to June 1967., J. Geophys. Res., 75, 4643-4657, 1970.
15. BURLAGA, L.F. and OGILVIE, K.W., Heating of the solar wind, Astrophys. J., 159, 1970, 659-670.
16. LEMAIRE, J., *Unpublished material*, 1988.
17. PANNEKOEK, A., Ionization in stellar atmospheres, bull. Astron. Inst. Netherlands, 19, 1922, 107-118.
18. ROSSELAND, S., Electrical state of a star, Month. Notices Roy. Astron. Soc., 84, 1924, 720-728.
19. PILIPP, W.G., MIGGENRIEDER, H., MONTGOMERY, M.D., MUHLHAUSER, K.H., ROSENBAUER, H. and SCHWENN, R., Characteristics of electron velocity distribution functions in the solar wind derived from the Helios Plasma Experiment, J. Geophys. Res., 92, 1987, 1075-1092.
20. HAERENDEL, G., Microscopic plasma processes related to reconnection, J. Atmos. Terr. Phys., 40, 1978, 343-353.
21. LEMAIRE, J. et ROTH, M., Penetration of solar wind plasma elements into the magnetosphere, J. Atm. Terr. Phys., 40, 331-335, 1978.
22. LEMAIRE, J., RYCROFT, M.J. and ROTH, M., Control of impulsive penetration of solar wind irregularities into the magnetosphere by the interplanetary magnetic field direction, Planet. Space Sci., 27, 1979, 47-57.
23. LEMAIRE, J., Plasmoid motion across a tangential discontinuity (with application to the *magnetopause*), J. Plasma Phys., 33, 1985, 425-436.
24. SVALGAARD, L., Geomagnetic activity : dependence on solar wind parameters, in: Zirker, J.B. (ed.), Coronal Holes and High Speed Wind Streams, Colorado Associated University Press, 1977, 371-441.
25. CHUPP, E.L., High-energy particle production in solar flares (SEP, Gamma-Ray and Neutron Emissions), Physica Scripta, T18, 1987, 5-19.
26. LUNDIN, R., Processes in the magnetospheric boundary layer, Physica Scripta, T18, 1987, 85-102.
27. BAUMJOHANN, W. and PASCHMAN, G., Solar wind-magnetosphere coupling : processes and observations, Physica Scripta, T18, 1987, 61-72.
28. FÄLTHAMMAR, C.-G., Magnetosphere-ionosphere interactions - *Near-Earth manifestations of the* plasma universe, IEEE Trans. Plasma Science, PS-14, No.6, 1986, 616-628.
29. JOSELYN, J., Real-time prediction of global *geomagnetic activity*, in Proc. of Chapman Conf. on solar wind-magnetosphere coupling, 1986, 127-.
30. *MULDREW, D.B., Observations of the cusp with topside sounders*, in "The Polar Cusp", J. Holtet and A. Egeland (eds.), D. Reidel Publ. Co, Dordrecht, 1985, 377-386.
31. HEELIS, R.A., Interplanetary magnetic effects on high latitude ionospheric convection, in "The Polar Cusp", J. Holtet and A. Egeland (eds.), D. Reidel Publ. Co., Dordrecht, 1985, 293-304.
32. SANDHOLT, P.E., EGELAND, A., HOLTET, J.A., LYBEKK, B., SVENES, K., and ASHEIM, S., and DEEHR, C.S., Large- and small-scale dynamics of the polar cusp, J. Geophys. Res., 90, 4407-4414, 1985.
33. SANDHOLT, P.E., DEEHR, C.S., EGELAND, A., LYBEKK, B., VIERECK, R. and ROMICK, G.J., Signatures in the dayside aurora of plasma transfer from the magnetosheath, J. Geophys. Res., 91, 1986, 10063-10079.
34. WINNINGHAM, J.D. and HEIKKILA, W.J., Polar cap auroral electron fluxes observed with ISIS I, J. Geophys. Res., 79, 1974, 949-957.
35. TITHERIDGE, J.E., Ionospheric heating beneath the magnetospheric cleft, J. Geophys. Res., 81, 1976, 3221-3226, 1976.
36. FRIEDMAN, H., The Sun's ionizing radiation, Physics of the Upper Atmosphere, Ch. 4, ed. J.A. Ratcliffe, *Academic Press*, New York, 1960.

37. GUSSENHOVEN, M.S., HARDY, D.A. and CAROVILLANO, R.L., Average electron precipitation in the polar cusps, cleft and cap, in "The Polar Cusp", J.A. Holtet and A. Egeland (eds.), D. Reidel Publ. Co, Dordrecht, 1985, 85-98.
38. JOCKERS, J., Solar wind models based on exospheric theory, Astron. Astrophys., 6, 1970, 219-239.
39. PILIPP, W.G., MIGGENRIEDER, H., MÜHLHÄUSER, K.-H., ROSENBAUER, H., SCHWENN, R. and NEUBAUER, F.M., Variations of electron distribution functions in the solar wind, J. Geophys. Res., 92, 1987, 1103-1118.
40. BOCHSLER, P., Solar wind ion composition, Physica Scripta, T18, 1987, 55-60.
41. LEMAIRE, J., Interpretation of Northward $B_z$ (NBZ) Birkeland current system and polar cap convection patterns in terms of the impulsive penetration model, in Magnetotail Physics, A.T.Y. Lui (ed.), Johns Hopkins University Press, Baltimore, 1987, 83-90.
42. HEIKKILA, W.J., Comment on electric field evidence on the viscous interaction at the magnetopause, by F.S. MOZER, Geophys. Res. Lett., 13, 1986, 233-234.
43. HEIKKILA, W.J., Inductive electric field at the magnetopause, Geophys. Res. Lett., 9, 1982, 877-880.
44. CHAPMAN, S. and FERRARO, V.C.A., The theory of the first phase of a geomagnetic storm, Terr. Mag. Atmos. Electr., 45, 245-268, 1940.
45. DEMIDENKO, I.I., LOMINO, N.S., PADALKA, V.G., RUTKEVICH, B.N., Plasma stream in an inhomogeneous transverse magnetic field, Sov. Phys. - Techn. Phys., 14, 16-22, 1969.

## 6. Acknowledgements

I wish to thank all my colleagues who have encouraged me directly or indirectly to pursue the study of solar wind-magnetosphere interaction. I am also much obliged to L. Bossy, P. Ledoux, and M. Nicolet for their valuable support during all these years. I wish also to say my great appreciation to M. Roth, W. Heikkila, C.-G. Fälthammar, R. Lundin, W. Baumjohann and D.S. Evans for stimulating discussions on these most exciting topics of space plasma physics. I have appreciated the administrative support I received from the "Ministère de l'Education Nationale et de la Culture Française", from the "Fonds National de la Recherche Scientifique", from the Director of the Institute of Aeronomy, and the help from the members of the personnel. I wish also to thank J. NISBET for reading the manuscript, and, for his useful comments.

## Appendix I

### Kinetic theory of solar wind acceleration in brief for slow and fast speed streams

As first pointed out by Chamberlain (Ref.11) the Coulomb collision mean free path of thermal coronal protons become larger than the characteristic density scale height, beyond a radial distance of 3 to 6 solar radii, depending on the energy of the colliding charged particles.

Beyond this radial distance, called the exobase level, the protons as well as the electrons can be considered to be collisionless (non-interacting ; at least in a zero-order approximation) : their Knudsen number is larger than unity in the coronal ion-exosphere above the exobase.

All electrons and protons with velocities larger than the gravitational escape velocity (275 km/sec at 5 $R_\odot$), can in principle evaporate out of the gravitational potential well, assuming that collisions can be neglected in the ion-exosphere. But since the thermal speed is $\sqrt{m_p/m_e}$ times larger for the electrons than for the protons, the net escape flux should be 42 times larger for the electrons than for the protons. This draws a huge negative polarization current out of the solar corona. The positive polarization charge density acquired by the corona sets up a charge separation electric field and an overall negative electrostatic potential difference between the exobase and infinity : the plasma becomes polarized as a result of the tendency for the electrons to escape at a higher rate than the heavier thermal ions.

The electric potential difference decelerates the electrons and reduces the number of them which are able to escape out of the gravitational plus electrostatic potential well. This electric potential difference accelerates the protons outward, thus increasing their net escape flux and their kinetic energy.

It has been shown (Ref.12, 38) that the equilibrium value of this electric potential difference adjusts itself to obtain escape fluxes which are equal for the retarded electrons and for the accelerated ions. Since, for the protons, the equilibrium value of the accelerating electrostatic force largely exceeds their decelerating gravitational force, all $H^+$-ions are drawn out of the corona and obtain supersonic velocities at large heliocentric distances, as indeed, is observed at 1 AU.

It has been indicated in Ref. 13 that Chamberlain also would have been able to predict the observed supersonic proton bulk speed at 1 AU, instead of a subsonic solar breeze, if he had used in 1961 the electric potential distribution described above, instead of using the standard Pannekoek (Ref.17) - Rosseland (Ref.18) electric field which is strictly applicable only when a plasma is in hydrostatic equilibrium in the gravitational field.

The exospheric potential difference for which the net electron flux balances the net ion escape flux is equal to - 700 Volts when exobase temperature of the electrons and protons are equal to $1.4 \times 10^6$K (Ref. 12). When the coronal exobase temperature is larger than this typical value a larger electrostatic potential sets up to maintain net zero electric current ; consequently for larger coronal exobase temperatures, larger proton bulk velocities are expected at 1 AU, as well as higher plasma temperatures (Ref. 12).

The positive correlation between solar wind temperatures and bulk velocities which is predicted by this simple kinetic theory fits nicely the observations (Refs. 14, 15) in low speed solar wind regions (i.e. where the measured bulk speed is smaller than 400 km/sec) (see Ref. 13).

In high speed solar wind streams, originating in low density, low temperature coronal holes, however, quite the opposite result has been found. This indicates that present solar wind theories and models, including the MHD or fluid models, fail to give the observed solar wind features in flow regimes where bulk speeds are larger than 400-500 km/s.

We have recently been able to show (Ref. 16), however, that the presence of an additional outward flux of suprathermal electrons, besides the thermal escape flux of low temperature coronal hole electrons, increases the equilibrium value of the negative electrostatic potential difference between the corona and infinity. This enhancement of the electric potential accelerates the exospheric protons up to a bulk velocity of 600-700 km/s.

The "strahl electrons" mainly observed in low density high speed solar wind regions (Ref. 24), constitutes an additional subpopulation of suprathermal electrons. We have calculated that the flux of suprathermal "strahl electrons" is indeed large enough compared to the escape flux of the coronal hole thermal electrons to increase the solar proton speed up to 600-700 km/s i.e. up to the values observed in high speed solar wind streams.

Similar high-speed solar wind velocities can only be obtained in modern MHD or multi-fluid models when ad hoc heating of the solar wind is assumed between the 2 and 10 $R_\odot$. It is not yet clear, however, where this additional energy should come from, and, how it could be dissipated at the right place in the coronal plasma.

Further details on this new explanation of fast speed solar wind streams will be published elsewhere.

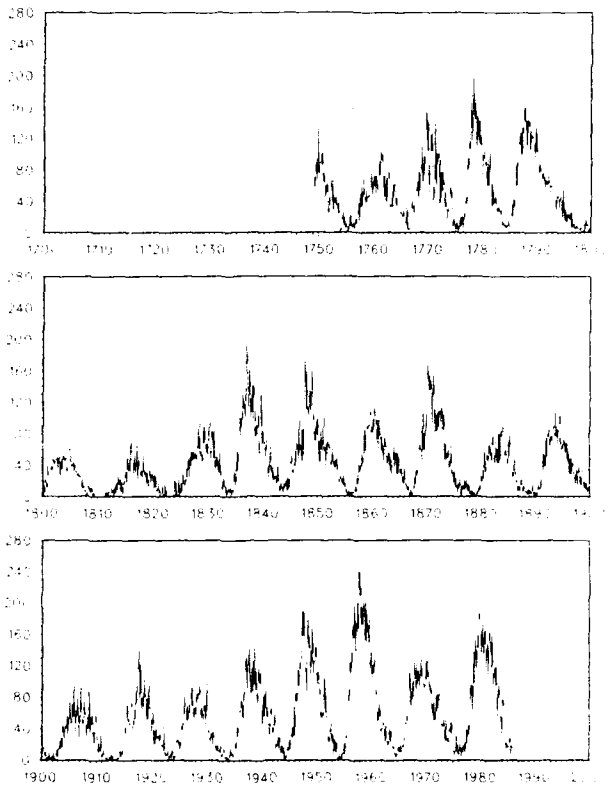



Fig.1 : MONTHLY MEAN SUNSPOT NUMBERS for the years 1700 - present time. Monthly mean numbers reveal the irregular cyclic variability. Monthly averages also contain a 27-day fluctuation that reflects the rotation period of the sun as the zonal distribution of spot groups sweeps across longitudes. The largest monthly mean value occurred in October 1947.

(from NGDC, Boulder)

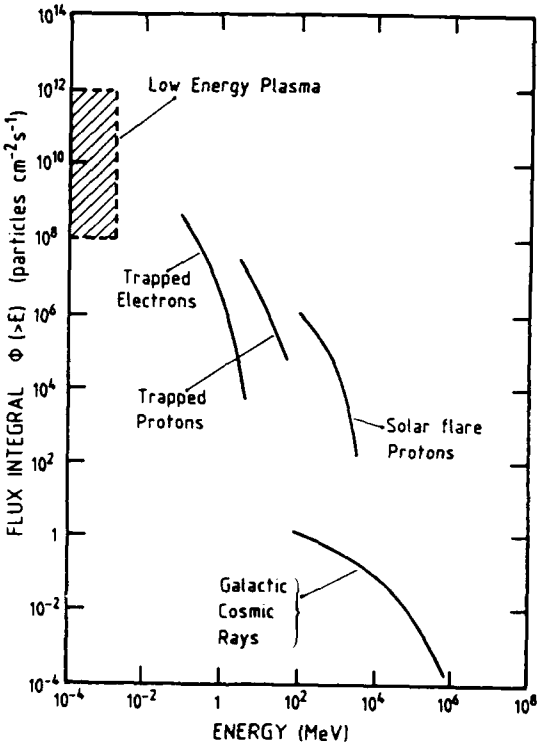



Figure 1 : Typical energy integral flux of the different populations of charged particles in the Earth magnetosphere.

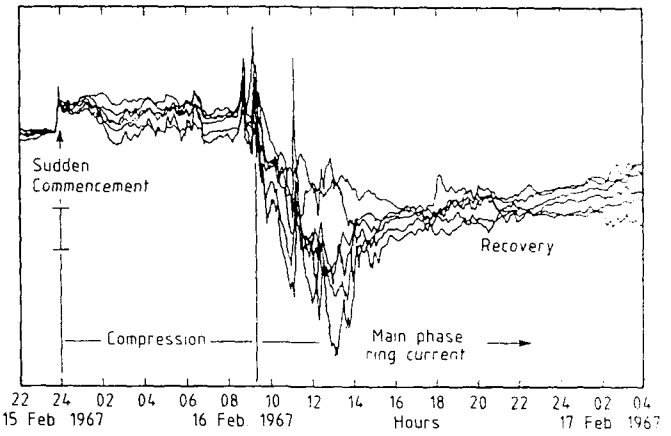



Fig.3 : A magnetic storm detected by the variation of north-south component of the magnetic field at six different magnetic observatories around the world. The "sudden commencement" and following sustained high field intensity results from the compression of the magnetosphere by an enhanced solar wind flow ; the pronounced decrease during the "main phase" is produced by the storm ring current that develops in the inner magnetosphere. Differences among stations are mainly due to longitudinal and local-time variations of the storm effects. (after Ref. 36)

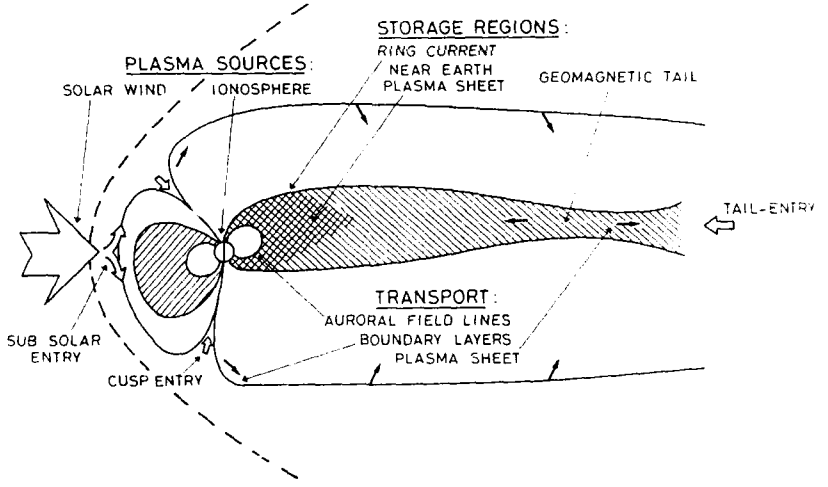



Fig.4 : meridional cross section of the Earth's magnetosphere showing different regions populated with charged particles of different energies and of different origins.

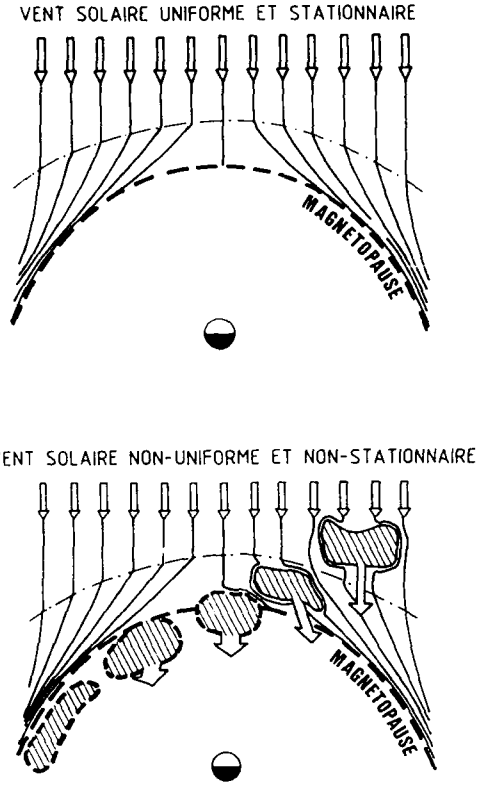



Fig. 1 : a) Equatorial cross section of a steady state magnetosphere with perfectly stationary and uniform solar wind flowing around it in a zero current and zero magnetospheric model.
b) Impulsive penetration of plasma irregularities present in solar wind plasma jet raining suddenly into the Earth magnetic field. As a result of their relative motion with respect to the magnetosheath and magnetosphere plasma, these raining plasmoids cause time-dependent (A.C.) electric and magnetic field perturbations, which cannot be described in the framework of DC electrostatic and magnetostatic theory and time-stationary field line distributions ! (after Ref. 21)

## DISCUSSION

**G.W.Prölss**

*Could you please comment on possible acceleration mechanisms in coronal holes?*

**Author's Reply**

Using the solar wind kinetic theory of *Lemaire and Scherer* [1971] one can roughly account for the bulk velocity and density observed in slow speed solar wind regions. By adding an escape flux of suprathermal "strahl" electrons to the escaping halo electrons one can increase the electrostatic potential between the corona and infinity, as indicated in the Appendix of my paper (in these proceedings); the larger electrostatic potential accelerates the ions to higher supersonic velocities like those observed in fast speed solar wind streams. The fact that the "strahl" electrons are almost exclusively observed in fast speed streams supports well this mechanism which I have proposed to explain the large solar wind velocities in fast speed streams. This explanation is based on simple kinetic theory. It does not involve extra (ad hoc) heat deposition of wave energy in the corona between 2 and 10 solar radii as would be needed in current hydrodynamical models to achieve the same high solar wind velocities at 1 AU.

# MAGNETOSPHERIC SUBSTORMS AS A SIGNATURE OF THE SOLAR TERRESTRIAL INTERACTION

Gordon Rostoker
Institute of Earth & Planetary Physics and
Department of Physics
University of Alberta
Edmonton, AB Canada T6G 2J1

Abstract: It is now well known that perturbations of the high latitude ionosphere are regulated by the properties of the solar wind and the interplanetary magnetic field. Variations in the direction of the interplanetary magnetic field lead to episodes normally lasting for a few minutes to a few hours in which large amounts of solar wind energy penetrate into the distant magnetosphere. This energy is stored in an intermediary form in the magnetotail magnetic field and in the kinetic drift energy of the earth's boundary layers and plasma sheet before it is ultimately degraded to heat through the action of electric current systems which couple the magnetosphere to the ionosphere. In this paper we shall discuss the physical processes through which the solar wind energy, once it has entered the magnetosphere, is redistributed and dissipated through the process known as the magnetospheric substorm. Based on the nature of the physical processes involved, we shall discuss the possibility of being able to predict the locale and level of the high latitude ionospheric perturbations which result from the action of the precipitation of energetic particles into the upper atmosphere as part of the electromagnetic coupling of the outer magnetosphere to the ionosphere.

## INTRODUCTION

It is now well recognized that solar wind particles are able to penetrate across the magnetopause and populate the boundary layers of the magnetosphere (Eastman et al., 1985)[1] as shown in Figure 1. The kinetic drift energy of the boundary layer plasma ultimately is redistributed through wave-particle interactions and through the action of magnetohydrodynamic (MHD) generator processes to heat a portion of the magnetospheric particle population and to drive the electric current systems which couple the magnetosphere to the ionosphere where energy is ultimately degraded to heat.

The entry of the solar wind plasma seems to be effected through two apparently different physical mechanisms. The first of these, termed the viscous interaction involves wave action at the magnetopause which results in the transfer of momentum (Miura, 1984)[2] or both momentum and mass across the magnetopause. This class of entry mechanism also is characteristic of the impulsive transport theories proposed by Heikkila (1982)[3] and by Lemaire (1985)[4]. The second class of entry mechanism has been termed magnetic field line reconnection (Vasyliunas, 1975)[5] and involves interconnection of solar wind and magnetospheric field lines. The merged field lines subsequently reconnect across the magnetotail neutral sheet (Coroniti, 1985)[6] and ultimately convect Earthward within the central plasma sheet (CPS). It is generally felt that the viscous interaction is responsible for the formation of the low latitude boundary layer (LLBL) found at the flanks of the magnetosphere centered at the midplane, while the reconnection is thought to be responsible for the formation of the plasma mantle located in the high latitude magnetotail between the magnetopause and the tail lobe. While there is at present no reason to believe that the viscous interaction is modulated by the direction of the interplanetary magnetic field (IMF), the reconnection rate clearly depends on the magnitude and direction of the IMF. Since the geomagnetic activity level in the magnetosphere has been demonstrated to depend on the polarity of the north-south (Bz) component of the IMF (Rostoker and Fälthammar, 1967)[7] in a manner predicted by reconnection theory (Dungey, 1961)[8], this has led to a general concensus that reconnection dominates the viscous interaction in terms of their respective contributions to the entry of the solar wind energy which is required to account for the observed energy dissipation in the auroral oval. In this respect, Baumjohann and Haerendel (1986)[9] suggest that reconnection processes are responsible for ~90% of the energy needed to power auroral processes. Further evidence for this approximate ratio between the contributions from reconnection processes and those from the viscous interaction come from the observations of Wygant et al. (1983)[10] who find that during episodes of northward interplanetary field, there is a constant cross polar cap potential drop of ~20 kV, while with increasing southward IMF there is an increasing cross polar cap potential drop ranging up to values of 100 - 150 kV (Figure 2). A discussion of the relative contributions of the reconnection process and the viscous interaction can be found in Rostoker (1987a)[11].

If one is to be able to predict episodes of time during which the high latitude ionosphere is disrupted insofar as transmission of information is concerned, one must have information about the plasma and field parameters of the interplanetary medium and how variations in those parameters influence the energy flow and dissipation in the magnetosphere-ionosphere system. Our understanding of the physics of this energy transport and dissipation comes largely from comparing the signatures of energy dissipation in the ionosphere to changes in the incoming solar wind. The signature of activity most often referred to nowadays in any serious correlation studies is the auroral electrojet or AE index. This index is obtained from the variations in the north-south component of the magnetic field perturbations measured at auroral zone sites distributed in a relatively uniform fashion around the northern hemisphere. Figure 3 shows a typical daily record of

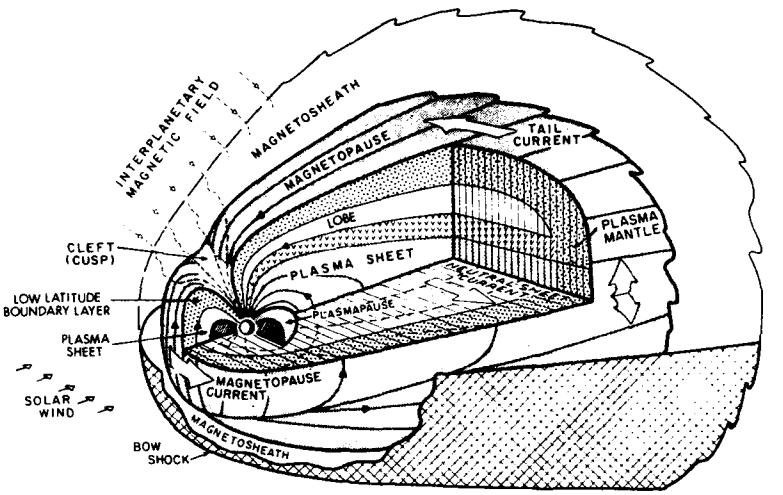



Figure 1 The particle and magnetic field regimes of the Earth's magnetosphere. The boundary layers are the site of anti-earthward flowing plasma which penetrates across the magnetopause from the magnetosheath while the plasma sheet is the site of earthward flowing warm plasma. The plasma sheet boundary layer at the high latitude edge of the (central) plasma sheet maps to the high latitude ionosphere where dynamic auroral activity is found.

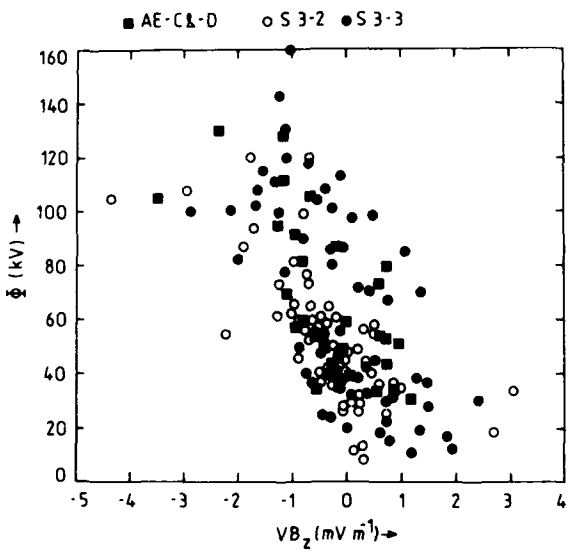



Figure 2 Response of the cross polar cap potential drop to the interplanetary electric field (after Cowley, 1984[12]). Episodes of northward IMF (positive $vB_z$) feature potential drops of 10 - 40 kV which don't vary noticeably with the magnitude of $vB_z$. When the IMF is southward (negative $vB_z$) the cross polar cap potential drop increases with increasing magnitude of $vB_z$.

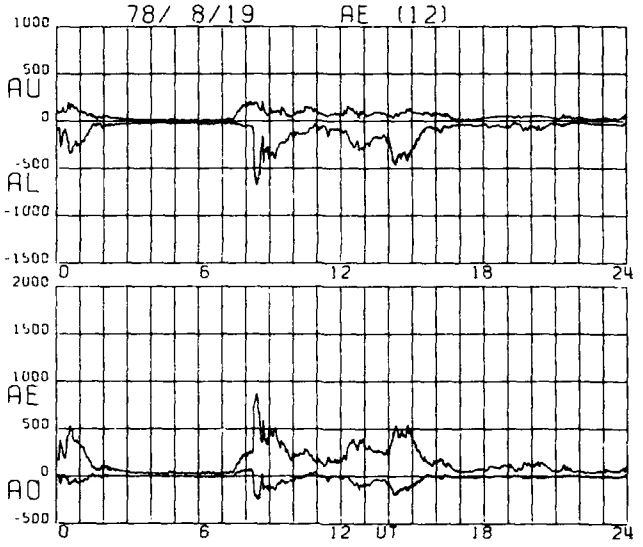



Figure 3 Record of the level of the auroral oval geomagnetic activity as quantified by the auroral electrojet (AE) index and the AU (maximum positive north-south component of the perturbation field) and AL (maximum negative north-south component of the perturbation field) indices which, together, make up AE. Substorm activity is clearly visible in the intervals 0000-0300 UT and 0730-1700 UT. The reader is referred to Rostoker (1972)[13] for the method of computation of the indices.

AE (together with the AU and AL indices which combine together to form AE). From this record, it is evident that fluctuations occur with many different time scales and over a large range of magnitudes. The largest perturbations seen on these records reflect a class of activity which has been given the name magnetospheric substorm. Anyone who wishes to study and predict ionospheric disruption at high latitudes ends up having to deal with the influence of substorm activity. Given its importance in the realm of energy dissipation in the ionosphere, it is perhaps surprising to realize that most researchers can't even agree on the definition of this disturbance. For that reason, before launching into a description of the physics of substorms, we must first provide a working definition of a substorm. In the next section we shall attempt to provide such a definition and to describe the characteristics of a substorm insofar as its ionospheric and auroral signatures are concerned.

## MAGNETOSPHERIC SUBSTORMS

One of the first manifestations of solar-terrestrial interaction which attracted the attention of scientists was unexpected deflections of compass needles at high latitude and their correlation with displays of brilliant and dynamic auroras as noted by A. Celsius in 1741. The magnetic perturbations associated with auroral outbursts as seen at a single auroral zone site were termed bays, and the perturbations could be seen in the north-south (H) component, the east-west (D) component or the vertical (Z) component of the perturbation field or simultaneously in any combination of these components. From the time of the International Geophysical Year around 1958, researchers began to piece together the patterns visible in the variability of magnetic signatures from station to station and in the patterns of auroral activity accompanying the magnetic variations. The first major breakthrough was the recognition that, at any instant, the auroras were arrayed around the world latitudinally confined belts centered near the geomagnetic poles (Feldstein, 1963)[14]. These belts were called the auroral ovals and were ultimately acknowledged to represent the projection on the ionosphere of the magnetotail plasma sheet (cf. Figure 1). For the next several years, attention focussed on relatively rapid spatially localized changes in the luminosity of the belts of auroras rather than on the belts in a global sense. Out of the studies in the early 1960's came the concept and definition of a substorm. The original term applied to auroral variations as introduced by Akasofu (1964)[15], and involved the brightening of a pre-existing discrete auroral arc in the midnight sector of the auroral oval. The disturbed region expanded poleward on a time scale of a few minutes and the western edge of the region developed a characteristic S-shaped structure. Because the western edge of the disturbed region expanded westward as the substorm evolved, this S-shaped structure was called the westward traveling surge.

The substorm evolution as proposed by Akasofu is shown in Figure 4.

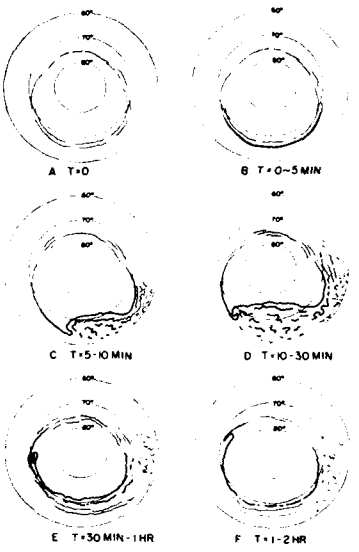



Figure 4 The evolution of an auroral substorm as deduced from ground based allsky camera data by Akasofu (1964)[15]. This synthesis of the data points to a sudden brightening of a pre-existing discrete auroral arc followed by explosive poleward expansion and formation of a single surge form which was thought to expand westward at velocities of ~1 km/s.

As studies of the influence of the interplanetary medium on auroral activity proceded, it became clear that southward turnings of the IMF led to the expansion of the polar cap area, manifested in the equatorward shift of the auroral oval. This expansion of the oval towards lower latitudes was normally accompanied by increasing ionospheric electrojet currents featuring eastward current across the dusk meridian and westward current across the dawn meridian. The rates of change of these large scale electrojet currents were typically much slower than for the sudden current changes which accompanied the auroral substorm of Akasofu (1964)[15] discussed above. However, the magnetic perturbations of the polar magnetic substorm (i.e. the electric current signature of the auroral substorm) and those of the larger scale eastward and westward electrojets are roughly of the same order ranging from tens of nT for weak activity to hundreds of nT for strong activity. A change in the Bz component of the IMF from northward to southward and back to northward again led to the magnetic activity level rising and then falling on the time scale of the change in the IMF. Since the IMF rises and falls occur on a rather unpredictable time scale ranging from minutes to hours, the magnetic perturbations (seen as magnetic bays in auroral zone magnetic records) have a considerable range of periods. Nonetheless, rises and falls in the magnetic activity level of duration 1 - 3 hrs have attracted the attention of many researchers and such perturbations have also been called "substorms". This has caused a problem in that when researchers use the term "substorm", one has to look carefully at exactly which set of magnetic and auroral signatures they are talking about.

This problem of notation was addressed first at a workshop in Victoria, Canada and published by Rostoker et al. (1980);[16] They agreed to give the term substorm to the normally longer time scale expansion and contraction of the auroral oval and used the terms "substorm expansive phase onset" and "substorm intensification" to describe the disturbance originally designated as a substorm by Akasofu (1964)[15]. In doing so, they acknowledged that many auroral breakups (as shown in Figure 4) could occur sequentially during the period of substorm development.

Further progress in understanding the nature of substorms came as the result of a workshop in Munster, FRG the results of which have been published recently by Rostoker et al. (1988)[17]. They noted that substorm activity appeared to involve two distinctive processes. In the directly driven process, energy from the solar wind enters the magnetosphere and is deposited directly in the auroral oval and ring current with delay times of only a few minutes. The auroral manifestation of the growth of the directly driven system is enhanced auroral luminosity on a global scale and the equatorward expansion of

the auroral oval. This is accompanied by strengthening of the eastward and westward electrojets. The decay of the directly driven system involves the poleward contraction of the auroral oval and the decay of the eastward and westward auroral electrojets. The loading-unloading process involves some of the solar wind energy being stored in the magnetic field and kinetic drift energy of the magnetotail field and particle configuration (i.e. loading). Subsequently, the energy is suddenly deposited in auroral oval and near-Earth ring current (i.e. unloading) with the amount of time during which the energy is stored being rather arbitrary and somewhat unpredictable at this time. Although it is known that northward turnings of the IMF can trigger episodes of unloading leading to substorm expansive phase activity (Rostoker, 1983)[18], it appears that some episodes of unloading may be triggered internally within the magnetosphere.

In summarizing the character of a magnetospheric substorm as it is presently envisioned, we can use the following description (which can be followed using Figure 5).

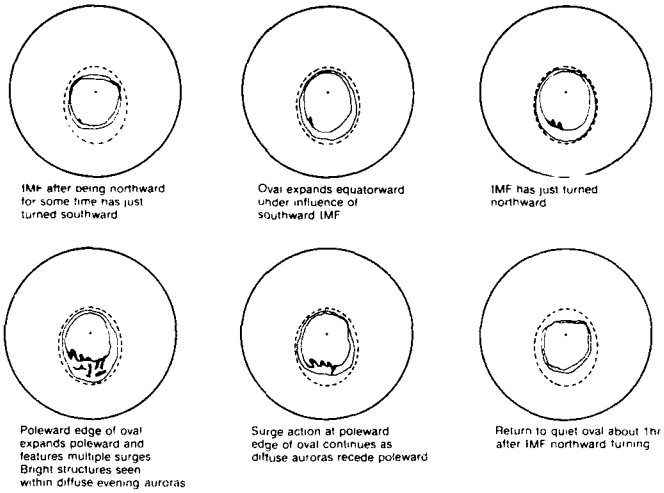



Figure 5 Evolution of auroral oval activity during the development of a substorm. The episode of energy input involves an expansion of the polar cap area (involving open field lines) with discrete auroral forms appearing episodically at the high latitude edge of the evening sector oval. The subsequent decline of energy input results in a decrease in the area of the polar cap and the excitation of significant auroral turbulence and surge like structures at the poleward edge of the oval as stored energy is deposited in the auroral ionosphere. The overall substorm magnetic disturbance stems from both temporal changes in directly driven electrojet intensity as well as poleward and equatorward motions of the electrojets superposed on which are the rapid transient changes of the multiple localized current wedges.

We begin by envisioning a situation where the IMF has been steady for the previous ~3 hours. We shall suppose that the north-south (Bz) component of the IMF is northward, although our scenario does not demand this polarity. We then consider a change in the IMF Bz component such that it becomes southward. (In principle the same sequence of events will happen if the Bz component becomes less northward but the effects will be accentuated if the Bz component actually attains a southward direction.) A consequence of this southward turning of the IMF is that the area of the polar cap (viz. the area of the polar ionosphere threaded by open field lines) begins to expand. This occurs through the equatorward shift of the poleward border of the auroral oval. As the polar cap increase in area proceeds, the driven system electrojets begin to increase in strength and, from time to time, small surge forms appear at the poleward edge of the evening sector auroral oval. While the surge forms may appear to be small scale, if one were to have a magnetometer at the poleward edge of the oval just to the east of the surge, a negative H-bay signature would be observed whose magnitude would be of the order of ~150 nT. However, the spatial extent of the causative westward electrojet is relatively small and as little as 100 - 200 km away from the electrojet current, the observed magnetic perturbation would be almost undetectable. In fact, the magnetic signature of the electric current associated with the small surge is precisely that of the substorm current wedge (cf. Baumjohann, 1983)[19]. However, the wedge-like signature of such a small high latitude surge would not be detectable at middle latitude stations. Over the period of growth of the area of the polar cap and driven system strength, small surges and their associated

current wedge signatures can occur sporadically at the poleward border of the evening sector oval.

Assuming that the southward IMF persists with relatively constant magnitude, one may use the impulse response time of the magnetosphere as a guide to understanding the ensuing behaviour of the auroral oval and its associated currents. Clauer et al. (1983)[20] have shown that input of solar wind energy to the magnetosphere can be represented by

$$O(T) = \int_{-\infty}^{\infty} g(t)\ I(t - \tau)d\tau$$

where g(t) is the impulse response of the magnetosphere. According to the studies of Clauer et al.(1983)[20] and Bargatze et al. (1985)[21], the impulse response of the magnetosphere has a time scale of ~2 hr. That is, if the magnetosphere in a specific state is affected by a change in IMF having the form of a Heaviside function, then the magnetosphere will attain a new and stable state after approximately 2 hr. Here we are alluding to the state of the directly driven system. Even during episodes of stability, it is still possible to observe sporadic surge activity at the poleward edge of the oval. It is not possible to say, at this time, whether this surge activity is due to external effects (such as small changes in solar wind conditions) or is due to internal triggering processes. Finally, it should be noted that the solar wind is normally rather turbulent with significant variations occurring on time scales normally much less than an hour. Thus, the magnetosphere rarely has an opportunity to move from one state to another in the manner predicted by the results of Clauer et al., since it is likely to be affected by a further change in solar wind conditions before it has a chance to reach a new steady state after a southward turning of the IMF.

Let us suppose, however, that after the southward turning of the IMF, the magnetosphere is exposed to constant interplanetary conditions for almost one complete substorm time constant (viz. ~1-2 hr.). Just before the interplanetary conditions change, the polar cap would have a larger area and the driven system electrojets would be stronger than before the original southward turning of the IMF. At this time, there might or might not be surge activity at the poleward edge of the evening oval. Now, let us suppose that the IMF suffers a sudden northward turning, returning the IMF configuration to its state prior to the original southward turning. At this time, the magnetosphere experiences a large enhancement in auroral and magnetic activity. In some cases, this activity is large enough to lead to significant mid latitude magnetic signatures, and many magnetospheric substorms described in the literature have their expansive phase marked from this time. At auroral oval latitudes, the consequence of the northward turning of the IMF is a rapid poleward motion of the poleward edge of the oval starting near midnight. As the poleward edge of the oval moves poleward, it is distorted by multiple surge forms stretching from midnight to the dusk meridian and sometimes even into the post-noon quadrant. The diffuse auroras south of the poleward border of the oval often acquire structure which may be aligned with convective and/or energetic electron drift paths (Rostoker et al., 1987)[22]. The multiple surge forms each have their associated current wedge producing a pattern of irregularly distributed intense westward electrojet elements distributed in the wake of the poleward advancing edge of the oval and along it's east-west extent in the evening sector. These electrojet elements are created in such rapid succession that new ones are created faster than the old ones can decay. Thus, the observed magnetic perturbation pattern comes from the ensemble of electrojet elements and their associated field aligned currents. Figure 6 shows the mid latitude and high latitude magnetic signatures across a wide range of longitudes during the course of a magnetosphere substorm. For this particularly large event, the poleward edge of the auroral oval was close to 66° N (geomagnetic). This leads to a situation in which many of the high latitude stations shown in Figure 6, which normally lie well within the auroral oval, find themselves either at the poleward edge of the oval or within the polar cap. These stations, while being poorly placed to detect the H-component signatures of the directly driven electrojets, are well placed to detect the sharp H-component signatures of the substorm expansive phase intensifications which are confined to the poleward border of the oval. The westward progression of the western edge of the substorm-disturbed region is clearly evident in the H-component signatures of the mid latitudes stations. (Here the diagnostic is the appearance of a positive H-component perturbation associated with passage of the upward field-aligned current at the western edge of the disturbed region across the meridian of the observing station.) During the height of the rapid succession of expansive phase intensifications described above, the poleward edge of the nightside oval contracts poleward as shown in Figure 5. Ultimately, if the IMF retains its northward polarity and strength, the expansive phase activity at the poleward border will subside and the diffuse auroras in the equatorward portion of the oval will gradually fade away. As the oval contracts, the directly driven electrojets decline in strength with characteristic decay times being of the order of 1/2 hr. to 2 hours. Ultimately, the oval will be reduced to a locus of weak discrete arcs at rather high latitudes and the equatorward edge of a weak and narrow region of diffuse auroras will establish itself just on the equatorward side of the discrete auroras.

In the foregoing section, I have described the phenomenology of substorm activity as seen in the auroral and magnetic activity. In the next portion of this paper, I would like to comment on the models which have been proposed to date which are designed to explain the physical processes which go on in the magnetotail during the evolution of substorm activity.

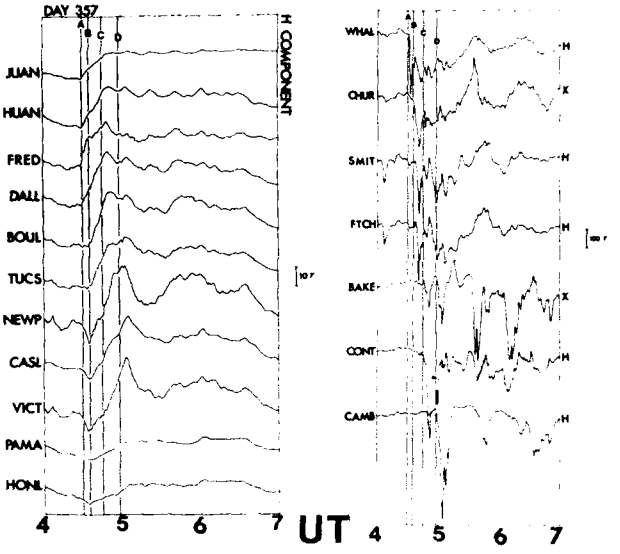



Figure 6 Magnetograms showing the north-south component of the magnetic perturbation field for stations stretching from east to west at middle latitudes (left panel) and for selected high latitude magnetograms at stations strongly influenced by the substorm seen in the figure. Lines A, B, C and D mark substorm onset and subsequent intensifications. The tendency for the disturbance to progress westward and to higher latitudes is evident. Each substorm intensification involves a poleward expansion of the region of discrete auroras and the formation of a surge in a longitudinally localized sector (adapted after Wiens and Rostoker, 1975[5]).

## MODELS FOR SUBSTORM ACTIVITY

At the present time, there are two models which have been proposed to explain how energy from the solar wind is redistributed and dissipated in the magnetosphere-ionosphere system through substorm activity. These two models involve rather different frameworks in which to explain the data although in the future ways may be found to reconcile the two approaches. The first is called the near-Earth Neutral Line model (hereinafter referred to as the NENL model) while the second has been named the boundary layer dynamics model (hereinafter referred to as the BLD model). The NENL model has evolved largely through the analysis of satellite data and intensive theoretical studies and numerical simulations of the plasma and field environment of the magnetotail. The reader is referred to Hones (1984)[6] for a comprehensive summary of the elements of the NENL model. The BLD model has developed largely through the detailed studies of ground based auroral signatures and magnetic field perturbations associated with substorm activity together with low altitude polar orbiting satellite measurements of the precipitating particles, electric fields and field-aligned currents at high latitudes. The reader is referred to Rostoker and Eastman (1987)[7] for a comprehensive summary of the BLD model. In the balance of this section, I shall attempt to outline the essential elements of these two competing models.

### (i) The Near-Earth Neutral Line Model

The elements of the NENL are best described by reference to Figure 7, which has been widely used for visualizing the general changes in magnetotail topology during the course of the evolution of a substorm. In this framework, the substorm begins with a southward turning of the IMF. Enhanced merging at the front side magnetopause is followed by transport of the newly opened field lines into the magnetotail which is accompanied by a thinning of the plasma sheet in the center of the tail. The growth of the tail magnetic field is consistent with a growth in the crosstail current and the edge effect of this tail current close to the Earth leads to the development of a region of very weak field somewhere in the region of 10 - 15 Re behind the Earth. At some time, reconnection of plasma sheet field lines begins in this weak field region however it is thought that this reconnection has no uniquely identifiable ionospheric signature. Ultimately, when all the (closed) plasma sheet field lines have been reconnected, reconnection of open lobe field lines commences. The tension in the reconnected field lines on the anti-sunward side of the near Earth neutral line leads to the ejection of the plasma sheet particles in a structure called a "plasmoid". The onset of reconnection of lobe field lines marks the onset of the substorm expansive phase and the substorm evolution proceeds with continuing

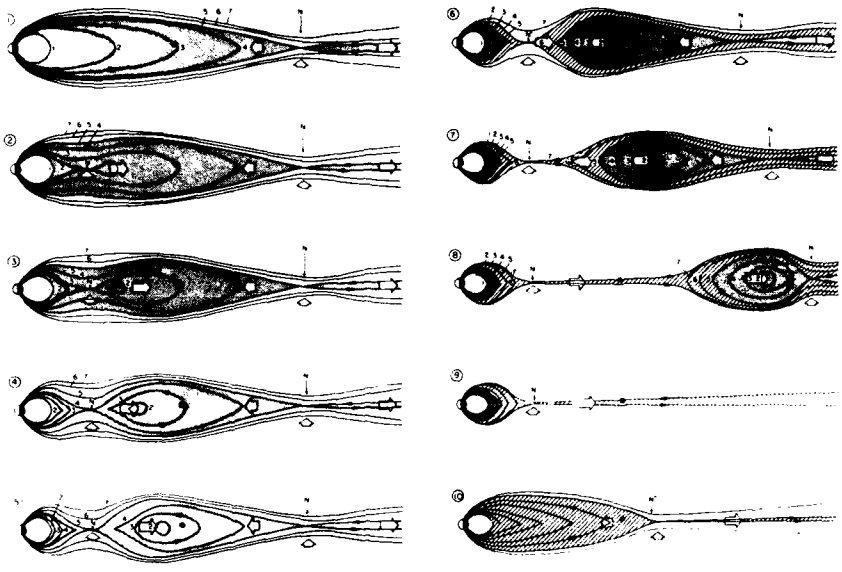


Figure 7 Evolution of a plasmoid in the near-Earth neutral line model for the magnetospheric substorm. The plasma sheet is pinched off close to the Earth and tension in the field lines on the anti-sunward side of the netural line causes a portion of the plasma sheet particle population to be returned to the interplanetary medium as a plasmoid moving downtail (after Hones, 1984[17]).

reconnection of lobe field lines. The neutral line stays at its location for some time and then, for reasons not presently understood, moves tailward accompanied by a thickening of the plasma sheet. In the ionosphere, the thickening of the plasma sheet is accompanied by a poleward motion of the auroras.

(ii) Boundary Layer Dynamics Model

The elements of the BLD model are best presented by reference to Figures 1 and 8. The essence of this model lies in the mapping to the magnetotail of the electric field polarity transition zones in the high latitude evening and morning sector high latitude auroral oval. In this scheme the velocity shear zones associated with the ionospheric electric field polarity transitions map through the plasma sheet boundary layer (PSBL) to the interface between the low latitude boundary layer (LLBL) and the central plasma sheet (CPS). In the solar terrestrial interaction, magnetosheath plasma penetrates the magnetopause through magnetic field merging (leading to the formation of the plasma mantle) and through action of a Kelvin-Helmholtz instability along the flanks of the magnetosphere (leading to the formation of the LLBL). Plasma mantle particles streaming downtail also convect towards the midplane of the tail where reconnection processes or current layer acceleration processes (cf. Lyons and Speiser, 1982)[27] lead to the redirection of some of the plasma mantle particles to form at least part of the Earthward convecting CPS particle population. The relatively steady processes are associated with the energetics of the directly driven process involving the large scale Birkeland current system (seen in Figure 8) and the eastward and westward auroral electrojets. Now one must look at the changes in velocity shear at the interface between the LLBL and CPS. In the conceptual framework where the tail is loaded during episodes of enhanced magnetic merging on the frontside magnetopause, magnetic field energy becomes available in the tail to power auroral processes (McPherron, 1970)[27]. Sporadic releases of this inductively stored energy (often stimulated by northward turnings of the IMF as demonstrated by Caan et al. (1977)[26] and by Rostoker (1983)[18]) can lead to bursts of reconnection and enhanced plasma flows in the plasma sheet. The Earthward flow on the Earthward side of the neutral line can enhance the momentum density in the CPS to the extent of making the LLBL/CPS interface unstable to the Kelvin Helmholtz instability. Processes consistent with Le Chatelier's Principle develop which attempt to stabilize the instability resulting in localized field-aligned current flow linking the high latitude ionosphere to the unstable shear zone in the magnetotail. These localized field-aligned currents are associated with

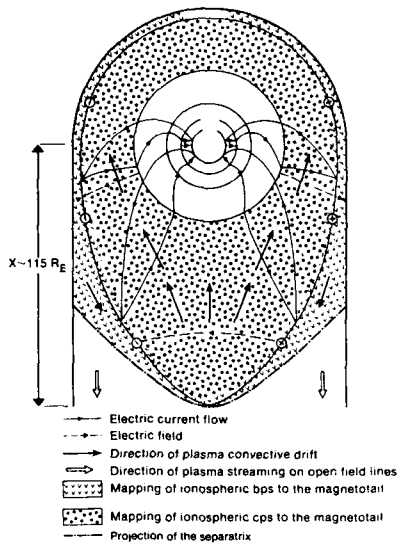



Figure 8 Projection on the plane of the neutral sheet of the particle populations of the magnetosphere together with electric field, electric current and convective flow directions in the LLBL and CPS. Dynamo processes lead to the extraction of kinetic flow energy from the CPS to drive a portion of the Birkeland current system (shown in the figure). The velocity shear zone between the LLBL and CPS maps to regions of active auroras in the high latitude portion of the auroral oval (after Rostoker and Eastman, 1987[25]).

substorm surges in the pre-midnight sector and to auroral omega bands in the dawn sector of the auroral oval (cf. Rostoker, 1987b).[29] Focussing our attention on the evening sector velocity shear zone (the ionospheric locus of which has been termed the Harang discontinuity) we associate the regions of localized upward field-aligned current with surface wave amplitude maxima along the LLBL/CPS interface in the magnetotail. The return currents flowing into the ionosphere around the core of upward field-aligned current (associated with closure of the electrical circuit) and linked to the upward currents by ionospheric Pedersen currents can provide the signature of the substorm current wedge as shown in Figure 9. Thus the ensemble of substorm current wedges which combine to create the overall substorm expansive phase perturbation can be thought of as a manifestation of the nonlinear growth of Kelvin-Helmholtz waves along the evening sector LLBL/CPS velocity shear zone as a consequence of bursts of enhanced reconnection at the neutral line. In summary, this model for magnetospheric substorm activity contains an explanation for both the background of directly driven activity as well as the more rapidly varying and spatially localized expansive phase activity.

## PREDICTION OF SUBSTORM ACTIVITY

We conclude this paper by commenting on our ability to predict high latitude ionospheric disruption through magnetospheric substorm activity in the light of our present understanding of the physical processes behind the substorm phenomenon. The comments on our ability to predict substorm activity are made in full realization of the fact that the most obvious and important indicator of the level of the solar-terrestrial interaction is the orientation of the incoming IMF. I shall break my comments about prediction capability into two sections. The first section will involve predictions when IMF and solar wind plasma data are available in near real time while the second will involve predictions when only ground based data are available.

### (i) Predictions Where Interplanetary Data Are Available

A considerable amount of work has been done in the past to find an algorithm involving the critical interplanetary plasma and field parameters which will permit the activity levels in the auroral regions to be predicted. Probably the most successful of these parameters is the parameter designed initially by Perrault and Akasofu (1978)[30] and having the form

$$\epsilon = \ell_0^2 \, v \, B^2 \sin^4 \frac{\theta}{2}$$

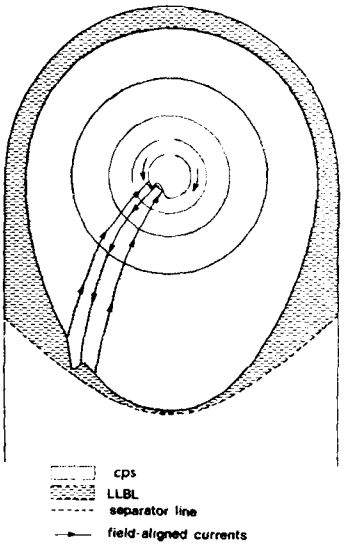



Figure 9 Schematic of current flow associated with an expansive phase (after de Groot-Hedlin and Rostoker, 1987). Growth of the Kelvin-Helmholtz instability at the LLBL/CPS interface due to enhanced Earthward CPS flow generated by a burst of reconnection causes localized current flow out of the ionosphere into the source region.

where v is the solar wind velocity, B is the magnitude of the interplanetary magnetic field, $\theta$ is the polar angle of the component of the IMF normal to the Sun-Earth line and measured from the northward geomagnetic axis and $l_o$ is a constant which has the dimensions of length and has a value of ~7 Re. The value of $\varepsilon$ is given in ergs/sec and ranges from about $10^{15}$ for low activity levels to ~$10^{19}$ for high activity levels. A physical justification for the precise analytical form of the $\varepsilon$ parameter can be found in Vasyliunas et al. (1982).

There is every indication that the $\varepsilon$ parameter is an excellent predictor of the strength and variation of the directly driven system involving the eastward and westward auroral electrojets (which are well quantified for average levels of magnetospheric activity by the AE index). Since the directly driven system accounts for a minimum of 50% of the energy dissipation involved in auroral oval activity, real time information regarding the IMF magnitude and direction together with the magnitude of the solar wind bulk flow velocity can go a long way in predicting auroral oval activity. Given the choice of IMF information or solar wind information, the IMF information is by far the most important as the solar wind velocity usually changes by a factor of two or less over extended periods of time whereas a simple polarity reversal of the IMF Bz component can change the parameter by as much as two to three orders of magnitude. It is also important to point out that substorm expansive phase activity is not related so simply to the $\varepsilon$ parameter. While the likelihood of expansive phase activity is enhanced with increasing $\varepsilon$, it is often observed that northward turnings of the IMF can trigger major substorm expansive phase activity (implying a possible anti-correlation of substorm expansive phase activity with the magnitude of $\varepsilon$). It would be fair to say that, in general all northward turnings of the IMF after some reasonable period of southward IMF (say a few minutes or more) result in substorm expansive phase activity with a delay of, at the most, a few minutes. However, it would appear that substorm expansive phase activity can also be initiated during periods when there is no clear northward turning of the IMF. A comprehensive study of the conditions for onset of substorm expansive phase activity still remains to be done and, until such a study is carried out, prediction of these large magnetospheric perturbations will still involve a significant amount of uncertainty.

(ii) Predictions Where Only Ground Based Data Are Available

Because of the poor existing coverage of interplanetary conditions and the lengthy amount of time which elapses before even that data which are recorded become available, anyone who wishes to predict magnetospheric activity with lead times of less than an hour is restricted to using data recorded on the ground using remote sensing techniques. During the International Magnetospheric Study (1976-1979), the modern technologies of communications and data management combined to make it possible for real time data to be obtained from arrays of ground instruments spread over large areas of the continental land masses. The one instrument useful in remote sensing magnetospheric conditions, the magnetometer, lent itself readily to the task of monitoring activity levels through detection of the magnetic perturbations from the large scale electric current systems which couple the magnetosphere to the ionosphere. Historically, the magnetometer provided the raw input data for computations of the Kp and AE indices which have long been used as a

quasi-quantitative measure of magnetospheric and ionospheric activity levels. In modern times, our knowledge of the forward model of these current systems permits us to monitor the level of the solar-terrestrial interaction to a certain extent. It is useful to discuss briefly the capabilities and limitations of ground magnetometer array data in predicting activity levels at auroral latitudes.

To begin with, there seems to be only one aspect of magnetospheric activity which is somewhat predictable using ground based data only. The predictability relates to the 27 day rotation period of the sun and involves the existence of active regions on the sun which maintain their integrity over as many as several solar rotation periods. These active regions appear to be the source of high speed solar wind flows and the impact on the magnetopause of these high speed streams in concert with a favourable orientation of the IMF leads to enhanced periods of magnetospheric activity which, in the past, have been termed M-region storms. Unfortunately, even for this rather crude predictor (crude because an active region on the sun can suddenly disappear in a quite unpredictable fashion) it turns out that it can only be used with confidence during periods near sunspot minimum. Figure 10 shows the anti-correlation of the 27 day recurrence with sunspot numbers. Sargent (1985)[13] has presented this relationship which shows that the 27 day recurrence breaks down as one moves into the ascending phase of the sunspot cycle. Thus one cannot safely count on using the 27 day recurrence for periods surrounding sunspot maximum.

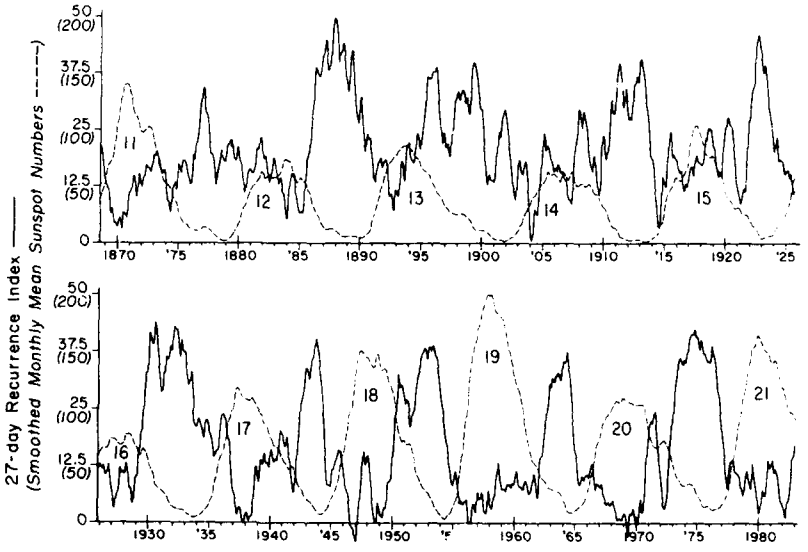



Figure 10 Sunspot cycle and 27-day recurrence index for the past century showing the tendency for recurrence to break down during periods of sunspot maximum (after Sargent, 1985[13]). This indicates that predictive capability breaks down just at the time when one expects activity levels to be the highest. Note that the 27-day recurrence pattern is the only repeatable feature of the solar terrestrial interaction which can be inferred from ground based data only. Its breakdown during periods of high sunspot numbers emphasizes the necessity of having real time interplanetary plasma and field parameters for effective prediction of magnetospheric substorm activity.

While it is difficult, if not impossible to predict activity levels well in advance using only ground based data, our improving knowledge of the detailed response of the auroral current systems to changes in the interplanetary medium now permits us to make some predictions about the probability of large substorm expansive phase disturbances (which can be directly linked to significant disruptions in communications where radio wave propagation is involved). In fact, it would be fair to say that two meridian lines of three component magnetometers separated in longitude by about 90°could be very effective in predicting substorm expansive phase activity of significant magnitude in the hour prior to its outbreak. The precise capabilities of such arrays involving stations distributed between approximately 55° and 90° (geomagnetic latitude) are outlined below:

(i) The ability of meridian lines such as those alluded to above to define the poleward edges of the eastward and westward directly driven auroral electrojets and hence the equatorward boundary of the polar cap. Knowledge of the area of the polar cap is equivalent to knowledge of the amount of magnetic energy stored in the magnetotail and hence of the amount of energy available for dissipation in the ionosphere in subsequent episodes of unloading.

(ii) The ability of meridian lines to define, through identification of the beginning of episodes of growth, when the interplanetary conditions have changed to enhance the flow of energy into the magnetosphere (e.g. a southward turning of the IMF). For example, in Figure 3, the episode of eastward electrojet growth identifiable between 0730 and 0815 UT is a clear indicator of enhanced energy input into the magnetosphere and the length of this episode of enhanced driven system activity in the absence of any significant expansive phase activity is a good indicator of the approach of a particularly large expansive phase outburst such as that seen at 0815 UT (which was triggered, as it turned out, by a sharp northward turning of the IMF).

One might also mention the fact that high pass filtered records from the same pair of station lines allows the use of ULF Pi 2 pulsations to define onsets of small scale substorm expansive phases which might not be so easily identifiable from normal magnetograms. In this respect, Pi 2 pulsations are similar to seismic fluctuations in terms of identifying impulsive outbursts of energy dissipation in the environment. Now that substorm researchers are able to distinquish between the physics of the directly driven process and the loading-unloading process, our ability to predict large scale disruptions of the auroral ionosphere are vastly improved compared to our capabilities in the past where the use of ground based data alone is involved. However, it is fair to say that only when we are in possession of real time interplanetary plasma and field data will prediction of magnetospheric activity be achieveable to a high degree of accuracy on time scales of minutes to hours.

ACKNOWLEDGEMENTS

This research was supported by the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

1. Eastman, T. E., L. A. Frank and C. Y. Huang, The boundary layers as the primary transport regions of the earth's magnetotail, J. Geophys. Res., 90, 9541-9560, 1985.
2. Miura, A., Anomalous transport by magnetohydrodynamic Kelvin-Helmholtz instabilities in the solar wind-magnetosphere interaction, J. Geophys. Res., 89, 801-818, 1984.
3. Heikkila, W. J., Impulsive plasma transport through the magnetopause, Geophys. Res. Lett., 9, 159, 1982.
4. Lemaire, J., Plasmoid motion across a tangential discontinuity (with application to the magnetopause), J. Plasma Phys., 33, 425, 1985.
5. Vasyliunas, V. M., Theoretical models of magnetic field line merging, 1., Rev. Geophys. Space Phys., 13, 303, 1975.
6. Coroniti, F. V., Explosive tail reconnection: the growth and expansive phases of magnetospheric substorms, J. Geophys. Res., 90, 7427-7448, 1985.
7. Rostoker, G. and C. -G. Fälthammar, Relationship between changes in the interplanetary magnetic field and variations in the magnetic field at the Earth's surface, J. Geophys. Res., 72, 5853, 1967.
8. Dungey, J. W., Interplanetary magnetic field and the auroral zones, Phys. Rev. Lett., 6, 47, 1961.
9. Baumjohann, W., and G. Haerendel, Dayside convection, viscous interaction and magnetic merging, in Solar Wind-Magnetosphere Coupling, Tokyo, Terra Scientific Pubi. Co., 1986.
10. Wygant, J. R., P. B. Torbert and F. S. Mozer, Comparison of S3-3 polar cap potential drops with the interplanetary magnetic field and models of magnetopause reconnection, J. Geophys. Res., 88, 5727-5735, 1983.
11. Rostoker, G., Dialog on the relative roles of reconnection and the "viscous" interaction in providing solar-wind energy to the magnetosphere, G. Rostoker, T. N. Baker, J. Lemaire and V. Vasyliunas) in Magnetotail Physics, Baltimore, MD, Johns Hopkins Univ. Press, 1987a, 409-414.
12. Cowley, S.W.-H., Solar wind control of magnetospheric convection, in Prog. Symp. Achievements of the IMS, ESA SP-217, 1984, 483.
13. Rostoker, G., Geomagnetic indices, Rev. Geophys. Space Phys. 10, 935, 1972.
14. Feldstein, Y. I., Some problems concerning the morphology of auroras and magnetic disturbances at high latitudes, Geomagnetizm i Aeronomiya 3, 183, 1963.
15. Akasofu, S. -I., The development of the auroral substorm, Planet. Space Sci. 12, 273-282, 1964.
16. Rostoker, G., S. -I. Akasofu, J. Foster, R. A. Greenwald, Y. Kamide, K. Kawasaki, A.T.Y. Lui, R. L. McPherron and C. T. Russell, Magnetospheric substorms - definition and signatures, Journal of Geophysical Research, 85, 1663-1668, 1980.
17. Rostoker, G., S. -I Akasofu, W. Baumjohann, Y. Kamide and R. L. McPherron, The roles of direct input of energy from the solar wind and unloading of stored magnetotail energy in driving magnetospheric substorms, Space Science Reviews, in press, 1988.
18. Rostoker, G., Triggering of expansive phase intensifications of magnetospheric substorms by northward turnings of the interplanetary magnetic field, J. Geophys. Res., 88, 6981-6993, 1983.
19. Baumjohann, W., Ionospheric and field-aligned current systems in the auroral zone: a concise review, in Advances in Space Research, 2, 55, 1983.
20. Clauer, C. R., R. L. McPherron and C. Searls, Solar wind control of the low-latitude asymmetric magnetic disturbance field, J. Geophys. Res., 83, 2123-2130, 1983.

21. Bargatze, L. F., D. N. Baker, R. L. McPherron and E. W. Hones, Jr., Magnetosphere impulse response for many levels of geomagnetic activity, J. Geophys. Res., 90, 6387-6394, 1985.
22. Rostoker, G., A.T.Y. Lui, C. D. Anger and J. S. Murphree, North-south structures in the midnight sector auroras as viewed by the VIKING imager, Geophys. Res. Lett., 14, 407-410, 1987.
23. Wiens, R. G. and G. Rostoker, Characteristics of the development of the westward electrojet during the expansive phase of magnetospheric substorms, J. Geophys. Res., 80, 2109-2128, 1975.
24. Perreault, P. and S. -I. Akasofu, A study of geomagnetic storms, Geophys. J. Roy. astron. Soc., 54, 547, 1978.
25. Hones, E. W., Jr., Plasma sheet behaviour during substorms, in Magnetic Reconnection in Space and Laboratory Plasmas, ed. by E. W. Hones, Jr., Washington, D. C., American Geophysical Geophys. 1984, *Monograph* 30.
26. Rostoker, G. and T. E. Eastman, A boundary layer model for magnetospheric substorms, J. Geophys. Res., 92, 12,187-12,201, 1987.
27. *Lyons, L. R. and T. W. Speiser, Evidence for current sheet acceleration in the* geomagnetic tail, J. Geophys. Res., 87, 2276-2286, 1982.
28. McPherron, R. L., Growth phase of magnetospheric substorms, J. Geophys. Res., 75, 5592, 1970.
29. Caan, M. N., R. L. McPherron and C. T. Russell, Characteristics of the association between the interplanetary magnetic field and substorms, J. Geophys. Res., 82, 4837-4842, 1977.
30. Rostoker, G., The Kelvin-Helmholtz instability and its role in the generation of the electric currents associated with Ps 6 and westward traveling surges, in Magnetotail Physics, A.T.Y. Lui, Baltimore, Johns Hopkins Univ. Press, 1987, 169-173.
31. De Groot-Hedlin, C. L. and G. Rostoker, Magnetic signatures of precursors to substorm expansive phase onset, J. Geophys. Res., 92, 5845-5846, 1987.
32. Vasyliunas, V. M., J. R. Kan, G. L. Siscoe and S. -I. Akasofu, Scaling relationships governing magnetospheric energy transfer, Planet. Space Sci., 30, 359-365, 1982.
33. Sargent III, H. H., Recurrent geomagnetic activity: evidence for long-lived stability in solar wind structure, J. Geophys. Res., 90, 1425-1428, 1985.

## DISCUSSION

**R.A.Hoffman**

You presented two models of a substorm, the deep tail model of plasmoids and the "ground-based" model. I do not understand that the models are irreconcilable. The plasmoid model breaks a body of magnetospheric plasma from the plasma sheet to move down the tail. After this occurs there is no communication from the plasmoid to the ionosphere. There is no comment regarding the plasma inside the region of reconnection. The ground-based model may address only the latter region, and its connection to the ionosphere. Your comments, please.

**Author's Reply**

First of all, the near earth neutral line model really makes no attempt to explain the detailed behavior of auroras and electric currents in the auroral oval at this time - perhaps it can be enhanced to provide such an explanation, but that remains to be seen. More important, it is now clear that a minimum of 50% of substorm energy deposition is associated with the directly driven system. There is presently no plan in the near earth neutral line model for an explanation of the directly driven system. However, the directly driven system is an integral part of the boundary layer dynamics model. Finally, I think the plasmoid should have some ionospheric signature. You have an azimuthally localized plasma blob sliding *rapidly* away from the earth adjacent to central plasma sheet material convecting earthward. There ought to be an ionospheric signature of this new and transient velocity shear zone. I know of no such transient velocity shear zones in the high latitude ionosphere.

**J.Lemaire**

I would like to mention that in addition to the 2 "predictions" of geomagnetic activity you showed, there is one which is much more detailed and which includes both of those you mentioned: this is the predictor determined by L. Svalgaard* in 1977. It also takes into account of the dependence of geomagnetic activity on the degree of inhomogeneity of the driving solar wind!

(*Geomagnetic Activity Dependence on Solar Wind Parameter: In *Coronal Holes and High Speed Wind Streams*, Ziker, J. B. ed, Colorado Associated Univ. Press, 1977, 371-441.)

**Author's Reply**

The influence of variability of the Interplanetary Magnetic Field (IMF) on the level of magnetospheric activity was well documented by Ballif and Jones back in the late 1960's and early 1970's. There certainly seems to be a statistically demonstrable trend for increased IMF variability to produced enhanced magnetospheric activity, as quantified by *Kp*.

My personal opinion is that a fluctuating IMF causes short term episodes of loading and unloading which leads to continual weak expansive phase activity. [The directly driven system is not much influenced by this irregularity of the IMF because of the long time scale ($\sim$ 2 hr) of the impulse response time of the magnetosphere.] I must emphasize that you can get larger levels of magnetospheric activity with quite low variance in the IMF, so such variability is not a requisite condition.

Finally, I would emphasize that there is a physical basis for the $\epsilon$ parameter [cf. *Vasyliunas et al.*, Planetary and Space Science, 1982] such that it represents energy input in terms of ergs $s^{-1}$. Involving the variance provides more of an empirical predictor, but it may well help create an empirical predictor, which is more effective that the present parameter in reproducing the overall energy input. A correct physical approach to developing a physically meaningful predictor demands separation of a substorm disturbance into directly driven and loading-unloading components.

# Balance equations for ionospheric plasmas with different partial temperatures

K. Suchy
Institute for Theoretical Physics, University of Duesseldorf
Universitaetsstr. 1, D-4 Duesseldorf 1, Fed. Rep. Germany

**Summary**

Different expressions for pressure tensors in multispecies plasmas are investigated for their physical meaning. Partial temperatures are introduced proportional to the partial translational energies (per particle). This leads to the definitions of translational energy tensors and temperature tensors proportional to the pressure tensors. Stress tensors are defined as *differences of the corresponding pressure tensors and their values in local* thermodynamic equilibrium. Heat flux vectors are defined as quantities with vanishing divergence under adiabatic constraints. Balance equations are established for masses, momenta, pressures, temperatures, energies, enthalpy, stresses and heat fluxes.

## 1. Introduction

In thermodynamic equilibrium the quantities pressure $p$, number density $n$, and temperature $T$ of an ideal gas are related by the equation of state

$$p = n k T \tag{1.1}$$

where $k$ is Boltzmann's constant. This law holds for a "simple" gas consisting of one species of particles. It must be dilute enough to neglect *the volume of the particles compared with the volume of the whole gas and* hot enough to neglect the interaction potentials of the particles compared with their translational kinetic energy. The latter two properties hold for almost all aeronomic plasmas. The first attribute, the one species property, can never hold for a plasma which, by its very definition, needs at least two species of (oppositely) charged particles. Occasionally a one-species approximation can be useful, but we want to concentrate on multi-species plasmas in this paper.

*An extension of the equation of state (1.1) to multi species ideal* gases is given by Dalton's law

$$p = \sum_j p_j = \sum_j n_j k T =: n k T \tag{1.2}$$

yielding the total pressure $p$ as the sum of the partial pressures $p_j$ of the different species $j$ with partial number densities $n_j$. The sum of the $n_j$ is the total number density $n$. This law holds also for aeronomic plasmas in thermodynamic equilibrium.

But thermodynamic equilibrium is a mathematically convenient limit which is very seldom reached for aeronomic plasmas. Most of them are in streaming motion, characterized by averaged velocities

$$\overline{\underset{\sim}{c}_j}(\underset{\sim}{r}, t) := \frac{1}{n_j(\underset{\sim}{r}, t)} \int d^3 c_j \, f_j(\underset{\sim}{c}_j, \underset{\sim}{r}, t) \, \underset{\sim}{c}_j \tag{1.3}$$

of the different species. The velocity distribution functions $f_j(\underset{\sim}{c}_j, \underset{\sim}{r}, t)$ are normalized as

$$n_j(\underset{\sim}{r}, t) = \int d^3 c_j \, f_j(\underset{\sim}{c}_j, \underset{\sim}{r}, t). \tag{1.4}$$

Averages over the velocity spaces are denoted by overbars, as $\overline{\underset{\sim}{c}_j}$ in (1.3).

Several questions immediately arise:

a) Are there still relationships

$$p_j = n_j k T$$

between the partial pressures $p_j$ and a quantity "temperature" $T$ common for all particle species?

b) Does it make sense to introduce partial temperatures $T_j$ and how are they related to $p_j$ and $T$ ?

c) Is Dalton's law (1.2) still valid?

d) Are scalar quantities $p_j$, $p$ still sufficient to describe partial pressures and the total pressure?

In the following chapters 2 and 3 these questions are answered, starting with the last question. This leads to physically meaningful definitions of pressures, temperatures, and stresses. For pressures, temperatures and some other quantities balance equations are established in chapter 4, leading to a physically resonable definition of heat flux vectors. For these and for the stresses balance equations are derived in chapter 5.

## 2. Pressures and translational energies

To look for physically resonable definitions of partial pressures and the total pressure we need to describe the transport of the momenta $m_j \underset{\sim}{c}_j$ across test surfaces $\hat{\underset{\sim}{a}}_j$ (of unit area). For later use we generalize this, looking for the transport of any velocity dependent quantity $\phi_j(\underset{\sim}{c}_j)$.

For a non-streaming plasma (with all $\overline{\underset{\sim}{c}_j} = 0$ ) and test surfaces $\hat{\underset{\sim}{a}}_j$ at rest the $\phi_j$ -fluxes across $\hat{\underset{\sim}{a}}_j$ are

$$\hat{\underset{\sim}{a}}_j \cdot \overline{\underset{\sim}{c}_j \, n_j \, \phi_j(\underset{\sim}{c}_j)}$$

since $\hat{\underset{\sim}{a}}_j \cdot \underset{\sim}{c}_j$ is the volume per unit time out of which particles with $\underset{\sim}{c}_j$ cross the test surface $\hat{\underset{\sim}{a}}_j$ . For a streaming plasma with test surfaces $\hat{\underset{\sim}{a}}_j$ moving with velocities $\underset{\sim}{u}_j$ these $\phi_j$ -fluxes become

$$\hat{\underset{\sim}{a}}_j \cdot \overline{(\underset{\sim}{c}_j - \underset{\sim}{u}_j) \, n_j \, \phi_j(\underset{\sim}{c}_j - \underset{\sim}{u}_j)} . \tag{2.1}$$

With the mass densities

$$\rho_j := m_j n_j \tag{2.2}$$

the momentum fluxes

$$\begin{aligned} &\hat{\underset{\sim}{a}}_j \cdot \overline{(\underset{\sim}{c}_j - \underset{\sim}{u}_j) \, n_j \, m_j (\underset{\sim}{c}_j - \underset{\sim}{u}_j)} \\ &= \hat{\underset{\sim}{a}}_j \cdot \overline{(\underset{\sim}{c}_j - \underset{\sim}{u}_j) \, \rho_j \, \underset{\sim}{c}_j} + \hat{\underset{\sim}{a}}_j \cdot \overline{(\underset{\sim}{c}_j - \underset{\sim}{u}_j)} \, \rho_j \, \underset{\sim}{u}_j \end{aligned} \tag{2.3}$$

are the pressures of $j$ particles upon moving test surfaces $\hat{\underset{\sim}{a}}_j$ . If these test surfaces reflect the particles elastically, then the averaged normal velocities vanish:

$$\hat{\underset{\sim}{a}}_j \cdot \overline{(\underset{\sim}{c}_j - \underset{\sim}{u}_j)} = 0 . \tag{2.4}$$

The factors of the (unit) test surfaces $\hat{\underset{\sim}{a}}_j$ in (2.3) are the pressure tensors

$$\overline{(\underset{\sim}{c}_j - \underset{\sim}{u}_j) \, \rho_j \, \underset{\sim}{c}_j} = \rho_j \overline{\underset{\sim}{c}_j \, \underset{\sim}{c}_j} - \underset{\sim}{u}_j \, \rho_j \, \overline{\underset{\sim}{c}_j} , \tag{2.5}$$

depending on the velocities $\underset{\sim}{u}_j$ of the test surfaces $\hat{\underset{\sim}{a}}_j$ . A conservation of the partial (averaged) angular momenta would require the symmetry of the pressure tensors [1]. Therefore the velocities $\underset{\sim}{u}_j$ of the test surfaces $\hat{\underset{\sim}{a}}_j$ should be equal to the averaged velocities $\overline{\underset{\sim}{c}_j}$ (1.3) which are all different in general. Therefore each test surface $\hat{\underset{\sim}{a}}_j$ would be required to be sensitive only to the particles of species $j$ and we would end up with osmotic pressure tensors

$$\underset{\sim}{\sigma}_j := \rho_j \left( \overline{\underset{\sim}{c}_j \, \underset{\sim}{c}_j} - \overline{\underset{\sim}{c}_j} \, \overline{\underset{\sim}{c}_j} \right) = \rho_j \overline{(\underset{\sim}{c}_j - \overline{\underset{\sim}{c}_j})(\underset{\sim}{c}_j - \overline{\underset{\sim}{c}_j})} . \tag{2.6}$$

Now these semi-permeable test surfaces $\hat{\underset{\sim}{a}}_j$ are physically unrealistic in plasmas and the partial angular momenta are not conserved because of the interactions between particles of different species. But the total angular momentum is conserved. We therefore require a realistic test surface $\hat{\underset{\sim}{a}}$ reflecting elastically particles of all species. If $\underset{\sim}{u}$ is the velocity of $\hat{\underset{\sim}{a}}$ we can describe the sum of the momentum fluxes of all species as

$$\hat{\underset{\sim}{a}} \cdot \sum_j \overline{(\underset{\sim}{c}_j - \underset{\sim}{u})\, \rho_j\, (\underset{\sim}{c}_j - \underset{\sim}{u})} = \hat{\underset{\sim}{a}} \cdot \sum_j \overline{(\underset{\sim}{c}_j - \underset{\sim}{u})\, \rho_j\, \underset{\sim}{c}_j} - \sum_j \hat{\underset{\sim}{a}} \cdot \overline{(\underset{\sim}{c}_j - \underset{\sim}{u})\, \rho_j}\, \underset{\sim}{u} \,. \tag{2.7}$$

This is the total pressure of the plasma upon the test surface $\hat{\underset{\sim}{a}}$. Assuming elastic reflection of all particles, i.e.

$$\hat{\underset{\sim}{a}} \cdot \overline{(\underset{\sim}{c}_j - \underset{\sim}{u})} = 0 \tag{2.8}$$

we are left with the pressure tensor

$$\sum_j \overline{(\underset{\sim}{c}_j - \underset{\sim}{u})\, \rho_j\, \underset{\sim}{c}_j} = \sum_j \rho_j\, \underline{\underline{\overline{\underset{\sim}{c}_j\, \underset{\sim}{c}_j}}} - \underset{\sim}{u} \sum_j \rho_j\, \overline{\underset{\sim}{c}_j} \tag{2.9}$$

It must be symmetric because of the conservation of the total angular momentum (for interaction forces in the directions between the particles). Thus the velocity $\underset{\sim}{u}$ of the test surface $\hat{\underset{\sim}{a}}$ equals the mean mass velocity

$$\underset{\sim}{v} := \frac{1}{\rho} \sum_j \rho_j\, \overline{\underset{\sim}{c}_j} \qquad \left(\text{with } \rho := \sum_j \rho_j\right) \tag{2.10}$$

of the whole plasma. The total pressure tensor

$$\underline{\underline{p}} := \sum_j \rho_j\, \underline{\underline{\overline{\underset{\sim}{c}_j\, \underset{\sim}{c}_j}}} - \sum_j \rho_j\, \underset{\sim}{v}\, \underset{\sim}{v} = \sum_j \rho_j\, \overline{(\underset{\sim}{c}_j - \underset{\sim}{v})(\underset{\sim}{c}_j - \underset{\sim}{v})} \tag{2.11}$$

is the sum of partial pressure tensors

$$\underline{\underline{p}}_j := \rho_j\, \overline{(\underset{\sim}{c}_j - \underset{\sim}{v})(\underset{\sim}{c}_j - \underset{\sim}{v})} = \underline{\underline{\sigma}}_j + \rho_j\, (\overline{\underset{\sim}{c}_j} - \underset{\sim}{v})(\overline{\underset{\sim}{c}_j} - \underset{\sim}{v}) \tag{2.12}$$

Therefore Dalton's law

$$\underline{\underline{p}} = \sum_j \underline{\underline{p}}_j = \sum_j \underline{\underline{\sigma}}_j + \sum_j \rho_j\, (\overline{\underset{\sim}{c}_j} - \underset{\sim}{v})(\overline{\underset{\sim}{c}_j} - \underset{\sim}{v}) \tag{2.13}$$

holds for the partial pressure tensors $\underline{\underline{p}}_j$ (defined with a physically realistic common test surface $\hat{\underset{\sim}{a}}$), but not for the osmotic pressures $\underline{\underline{\sigma}}_j$ (2.6).

Because the pressure tensors $\underline{\underline{p}}_j$, $\underline{\underline{p}}$ are real and symmetric, their eigenvalues $p_j^\alpha$, $p^\alpha$ ($\alpha = 1, 2, 3$) are real and their eigenvectors are real and orthogonal. With the peculiar velocities

$$\underset{\sim}{C}_j := \underset{\sim}{c}_j - \underset{\sim}{v} \tag{2.14}$$

and their components $C_j^\alpha$ in the directions of the three eigenvectors, the eigenvalues $p_j^\alpha$ are expressed as

$$\frac{1}{2}\, p_j^\alpha = n_j\, \frac{m_j}{2}\, \overline{(C_j^\alpha)^2} =: u_j^\alpha \,. \tag{2.15}$$

These are the translational energy densities for particles $j$ moving only in the direction of the eigenvectors (as seen in a frame moving with the mean mass velocity $\underset{\sim}{v}$). They are the eigenvalues of the partial translational energy density tensors

$$\underline{\underline{u}}_j := \frac{1}{2}\, \underline{\underline{p}}_j \tag{2.16}$$

which can be summed up to the total translational energy tensor

$$\underset{\approx}{u} := \sum_j \underset{\approx}{u}_j = \frac{1}{2}\,\underset{\approx}{p} \,. \tag{2.17}$$

Having anwered the two questions regarding pressures we turn now to the temperatures.

## 3. Temperatures, hydrostatic pressures and stresses

In thermodynamic equilibrium the averaged energy per particle per degree of freedom equals $kT/2$ . Generalizing this we put

$$\frac{m_j}{2}\overline{\left(C_j^{\alpha}\right)^2} = \frac{u_j^{\alpha}}{n_j} =: \frac{k T_j^{\alpha}}{2} \,, \tag{3.1}$$

i.e. the averaged translational energy for particles $j$ moving in the direction of a pressure eigenvector is proportional to the corresponding eigenvalue $T_j^{\alpha}$ of a partial temperature tensor

$$\underset{\approx}{T}_j := \frac{1}{k}\, m_j \overline{\underset{\sim}{C}_j \underset{\sim}{C}_j} = \frac{1}{n_j k}\,\underset{\approx}{p}_j = \frac{2}{n_j k}\,\underset{\approx}{u}_j \,. \tag{3.2}$$

cf. (2.12) (2.14) (2.16). Thus we have generalized the equation of state $p = n k T$ (1.2) to tensor equations for partial quantities

$$\underset{\approx}{p}_j = n_j k \underset{\approx}{T}_j \,. \tag{3.3}$$

With the generalized Dalton law (2.13) we define the mean temperature tensor as

$$\underset{\approx}{T} := \frac{1}{n k}\,\underset{\approx}{p} = \frac{2}{n k}\,\underset{\approx}{u} = \frac{1}{n}\sum_j n_j \underset{\approx}{T}_j \,. \tag{3.4}$$

For a plasma in a strong magnetic field one eigenvector of the pressure tensors is parallel to the magnetic field, the other two are perpendicular to it. The corresponding parallel and perpendicular eigenvalues are then related to those of the energy density and temperature tensors as

$$p_j^{\parallel} = 2 u_j^{\parallel} = n_j k T_j^{\parallel} \qquad\qquad p_j^{\perp} = 2 u_j^{\perp} = n_j k T_j^{\perp} \,. \tag{3.5}$$

For particles moving in an arbitrary direction the translational energy has three degrees of freedom and we put

$$\frac{m_j}{2}\sum_{\alpha=1}^{3}\overline{\left(C_j^{\alpha}\right)^2} = \frac{m_j}{2}\,\overline{C_j^2} = \frac{1}{n_j}\sum_{\alpha} u_j^{\alpha} =: \frac{u_j}{n_j} =: \frac{3 k T_j}{2} \tag{3.6}$$

thus defining partial translational energies $u_j$ and partial temperatures $T_j$ . These are related to the eigenvalues $u_j^{\alpha}$ , $T_j^{\alpha}$ and to the tensors $\underset{\approx}{u}_j$ , $\underset{\approx}{T}_j$ and $\underset{\approx}{p}_j$ by

$$\begin{aligned} T_j &= \frac{1}{3}\sum_j T_j^{\alpha} = \frac{2}{3 n_j k}\sum_{\alpha} u_j^{\alpha} = \frac{2 u_j}{3 n_j k} \\ &= \frac{1}{3}\,\mathrm{trace}\,\underset{\approx}{T}_j = \frac{2}{3 n_j k}\,\mathrm{trace}\,\underset{\approx}{u}_j = \frac{1}{3 n_j k}\,\mathrm{trace}\,\underset{\approx}{p}_j \,. \end{aligned} \tag{3.7}$$

For a plasma in a strong magnetic field this means

$$T_j = \frac{1}{3}\left(T_j^{\parallel} + 2 T_j^{\perp}\right) \qquad\qquad u_j = u_j^{\parallel} + 2 u_j^{\perp} \tag{3.8}$$

since $T_j^{\perp}$ and $\omega_j^{\perp}$ are degenerate eigenvalues, each belonging to two eigenvectors perpendicular to the magnetic field and orthogonal but otherwise arbitrary.

Having defined partial translational temperatures $T_j$ as one third of the traces of the partial temperature tensors $\underset{\approx}{T}_j$ (3.2) we define the mean temperature $T$ with (3.4) as

$$T := \frac{1}{3} \operatorname{trace} \underset{\approx}{T} = \frac{1}{3nk} \operatorname{trace} \underset{\approx}{p} = \frac{1}{n} \sum_j n_j T_j \quad . \tag{3.9}$$

Analogously, with (2.17)

$$u := \operatorname{trace} \underset{\approx}{w} = \frac{1}{2} \operatorname{trace} \underset{\approx}{p} = n \frac{3kT}{2} \quad . \tag{3.10}$$

But instead of defining scalar partial pressures as one third of the traces of the partial pressure tensors $\underset{\approx}{p}_j$, which would lead to $n_j k T_j$ (3.7), we define partial "hydrostatic" pressures $p_j$ as

$$p_j := n_j k T \, . \tag{3.11}$$

The reason is the following: Without thermodynamic forces, like gradients of $\underset{\sim}{v}$ and $T$, the plasma is in local thermodynamic equilibrium. In this state the plasma is merely characterized by the parameters $n_j(\underset{\sim}{r}, t)$, $\underset{\sim}{v}(\underset{\sim}{r}, t)$ and $T(\underset{\sim}{r}, t)$ with their local and temporal values at $(\underset{\sim}{r}, t)$. These few parameters describe the plasma at $\underset{\sim}{r}, t$ in a way as if it were homogeneous and stationary with the same parameters $n_j$, $\underset{\sim}{v}$, $T$ at each position $\underset{\sim}{r}$ and at any time $t$. *In this case* the partial pressure tensors (2.12) are isotropic, i.e. proportional to the unit tensor $\underset{\approx}{I}$ as $n_j k T \underset{\approx}{I}$, and the total pressure tensor $\underset{\approx}{p}$ (2.11) is $nkT \underset{\approx}{I}$. Different partial temperatures $T_j$ (3.7), even if physically existing during a finite time interval in a finite domain, could not be maintained during all times, since different partial translational energies $3kT_j/2$ (3.6) are equalized, if no thermodynamic forces exist.

Therefore we have defined the partial hydrostatic pressures as $p_j := n_j k T$ (3.11) and define further the total hydrostatic pressure

$$p := nkT = \sum_j p_j = \frac{1}{3} \operatorname{trace} \underset{\approx}{p} \, . \tag{3.12}$$

The deviations

$$\underset{\approx}{\pi}_j := \underset{\approx}{p}_j - p_j \underset{\approx}{I} = \underset{\approx}{p}_j - n_j k T \underset{\approx}{I} = n_j \left( m_j \overline{\underset{\sim}{C}_j \underset{\sim}{C}_j} - kT \underset{\approx}{I} \right) \tag{3.13}$$

and

$$\underset{\approx}{\pi} := \underset{\approx}{p} - p \underset{\approx}{I} = \sum_j \underset{\approx}{\pi}_j \tag{3.14}$$

from the values in local thermodynamic equilibrium are called stress tensors in plasma physics.

The traces of the partial stress tensors $\underset{\approx}{\pi}_j$, viz.

$$\operatorname{trace} \underset{\approx}{\pi}_j = \operatorname{trace} \underset{\approx}{p}_j - 3 p_j = 3 n k (T_j - T) \tag{3.15}$$

represent the deviations of the partial temperatures $T_j$ (3.7) from the mean temperature $T$ (3.9). Their sum vanishes because of the definition (3.9) of $T$, thus leaving the total stress tensor $\underset{\approx}{\pi}$ (3.14) traceless:

$$\operatorname{trace} \underset{\approx}{\pi} = 0 \, . \tag{3.16}$$

Now all *questions* from the introduction are answered and we start to establish balance equations for pressures, temperatures, energies and also for masses, momenta and enthalpy in the next chapter.

4. Balances for masses, momenta, pressures, temperatures, energies, enthalpy.

Balance equations for quantities $\phi_j(\underset{\sim}{C}_j, \underset{\sim}{r}, t)$ depending on the peculiar velocities $\underset{\sim}{C}_j := \underset{\sim}{c}_j - \underset{\sim}{v}(\underset{\sim}{r}, t)$ (2.14) and additional macroscopic parameters varying with $(\underset{\sim}{r}, t)$ read [2]

$$\frac{\partial}{\partial t} n_j \overline{\phi_j} + \underset{\sim}{\nabla} \cdot \left( \underset{\sim}{v}\, n_j \overline{\phi_j} + n_j \overline{\underset{\sim}{C}_j \phi_j} \right) - n_j \left( \overline{\underset{\sim}{C}_j \cdot \underset{\sim}{\nabla} \phi_j} + \frac{D\overline{\phi_j}}{Dt} \right)$$

$$+ n_j \overline{\underset{\sim}{C}_j \cdot (\underset{\sim}{\nabla}\underset{\sim}{v}) \cdot \frac{\partial}{\partial \underset{\sim}{C}_j} \phi_j} + n_j \left( \frac{D\underset{\sim}{v}}{Dt} - \overline{\dot{\underset{\sim}{c}}_j} \right) \cdot \overline{\frac{\partial}{\partial \underset{\sim}{C}_j} \phi_j} = n_j \overline{\left( \frac{\delta \phi_j}{\delta t} \right)^{coll}} \tag{4.1}$$

with the convective derivative

$$\frac{D}{Dt} := \frac{\partial}{\partial t} + \underset{\sim}{v} \cdot \underset{\sim}{\nabla} \tag{4.2}$$

and the external forces $m_j \dot{\underset{\sim}{c}}_j$. If the latter contain Lorentz and/or Coriolis forces, thus depend on $\underset{\sim}{C}_j$, we decompose

$$\dot{\underset{\sim}{c}}_j = \dot{\underset{\sim}{c}}_{j0} + \underset{\sim}{\Omega}_j \times \underset{\sim}{C}_j \tag{4.3}$$

where $\dot{\underset{\sim}{c}}_{j0}$ is independent of $\underset{\sim}{C}_j$. The expressions on the right-hand sides of the balance equations (4.1) describe the collisional contributions. They are not evaluated in this paper.

The quantities

$$m_j \qquad m_j \underset{\sim}{C}_j \qquad m_j \underset{\sim}{C}_j \underset{\sim}{C}_j$$

do not depend on additional macroscopic parameters besides of $\underset{\sim}{v}(\underset{\sim}{r}, t)$. Therefore $\underset{\sim}{\nabla}\phi_j$ and $D\phi_j / Dt$ in the balance equations (4.1) vanish for these quantities and we obtain the mass balances

$$\frac{\partial \rho_j}{\partial t} + \underset{\sim}{\nabla} \cdot \overline{\underset{\sim}{c}_j \rho_j} = n_j \overline{\left( \frac{\delta m_j}{\delta t} \right)^{coll}}, \tag{4.4}$$

the momentum balances

$$\frac{\partial}{\partial t} \rho_j \overline{\underset{\sim}{C}_j} + \underset{\sim}{\nabla} \cdot \left( \underset{\sim}{v}\, \rho_j \overline{\underset{\sim}{C}_j} + \underset{\approx}{p}_j \right) + \rho_j \overline{\underset{\sim}{C}_j} \cdot \underset{\sim}{\nabla}\underset{\sim}{v}$$

$$+ \rho_j \left( \frac{D\underset{\sim}{v}}{Dt} - \dot{\underset{\sim}{c}}_{j0} \right) - \underset{\sim}{\Omega}_j \times \rho_j \overline{\underset{\sim}{C}_j} = n_j \overline{\left( \frac{\delta m_j \underset{\sim}{C}_j}{\delta t} \right)^{coll}}, \tag{4.5}$$

and the pressure (or energy) balances [3]

$$\frac{\partial}{\partial t} \underset{\approx}{p}_j + \underset{\sim}{\nabla} \cdot \left( \underset{\sim}{v}\, \underset{\approx}{p}_j + \rho_j \overline{\underset{\sim}{C}_j \underset{\sim}{C}_j \underset{\sim}{C}_j} \right) + 2 \widehat{\underset{\approx}{p}_j \cdot \underset{\sim}{\nabla}\underset{\sim}{v}}$$

$$+ 2 \widehat{\left( \frac{D\underset{\sim}{v}}{Dt} - \dot{\underset{\sim}{c}}_{j0} \right) \rho_j \overline{\underset{\sim}{C}_j}} - 2 \widehat{\underset{\sim}{\Omega}_j \times \underset{\approx}{p}_j} = n_j \overline{\left( \frac{\delta m_j \underset{\sim}{C}_j \underset{\sim}{C}_j}{\delta t} \right)^{coll}}. \tag{4.6}$$

A symmetric second order tensor $\widehat{\underset{\approx}{t}}$ is defined as half of the sum of $\underset{\approx}{t}$ and its transpose $\underset{\approx}{t}^T$, i.e.

$$2\widehat{\underset{\approx}{t}} := \underset{\approx}{t} + \underset{\approx}{t}^T. \tag{4.7}$$

Scalar multiplication of the pressure balances (4.6) from the left and from the right with the eigenvectors of $\underset{\sim}{p}_j$ yields balance equations for the eigenvalues $p_j^{\alpha} = 2u_j^{\alpha} = n_j k T_j^{\alpha}$ (2.15) (3.1), that means in strong magnetic fields for $p_j^{\parallel} = 2u_j^{\parallel} = n_j k T_j^{\parallel}$ and $p_j^{\perp} = 2u_j^{\perp} = n_j k T_j^{\perp}$. Using suitable combinations with the mass balances (4.4) one obtains balance equations for the temperature eigenvalues $T_j^{\alpha}$, i.e. for $T_j^{\parallel}$ and $T_j^{\perp}$ in strong magnetic fields.

Half the traces of the pressure balances (4.6) yield with (3.7) the partial translational energy balances

$$\frac{\partial}{\partial t} u_j + \underset{\sim}{\nabla} \cdot \left( \underset{\sim}{v}\, u_j + \frac{\rho_j}{2} \overline{\underset{\sim}{C}_j C_j^2} \right) + \underset{\approx}{p}_j : \underset{\sim}{\nabla}\, \underset{\sim}{v}$$

$$+ \left( \frac{D \underset{\sim}{v}}{Dt} - \dot{\underset{\sim}{c}}_{j0} \right) \cdot \rho_j \overline{\underset{\sim}{C}_j} = n_j \overline{\left( \frac{\delta}{\delta t} \frac{m_j}{2} C_j^2 \right)^{coll}} . \tag{4.8}$$

The sum over all particle species $j$, taking into account

$$\sum_j \rho_j \overline{\underset{\sim}{C}_j} = \sum_j \rho_j \overline{\underset{\sim}{c}_j} - \rho \underset{\sim}{v} = 0 \tag{4.9}$$

(because of the definitions (2.10) of $\underset{\sim}{v}$ and (2.14) of $\underset{\sim}{C}_j$) yields the total translational energy balance

$$\frac{\partial}{\partial t} u + \underset{\sim}{\nabla} \cdot \left( \underset{\sim}{v}\, u + \sum_j \frac{\rho_j}{2} \overline{\underset{\sim}{C}_j C_j^2} \right) + \underset{\approx}{p} : \underset{\sim}{\nabla}\, \underset{\sim}{v}$$

$$- \sum_j \dot{\underset{\sim}{c}}_{j0} \cdot \rho_j \overline{\underset{\sim}{C}_j} = \sum_j n_j \overline{\left( \frac{\delta}{\delta t} \frac{m_j}{2} C_j^2 \right)^{coll}} . \tag{4.10}$$

To derive a balance for the translational enthalpy per particle

$$h := \frac{5 k T}{2} = \frac{5 p}{2 n} \tag{4.11}$$

we decompose the total energy density $u$ with (3.10) (3.12) as

$$u = n \frac{3 k T}{2} = \frac{3 p}{2} = \frac{5 p}{2} - p$$

and the total pressure tensor $\underset{\approx}{p}$ with (3.14) as

$$\underset{\approx}{p} = p \underset{\approx}{I} + \underset{\approx}{\Pi} .$$

The terms with $p$ in the total translational energy balance (4.10) are then, with the mass balances (4.4),

$$\frac{\partial}{\partial t} \frac{5 p}{2} + \underset{\sim}{\nabla} \cdot \underset{\sim}{v}\, \frac{5 p}{2} - \frac{D p}{D t} = n \frac{D}{D t} \frac{5 p}{2 n} + \frac{5 p}{2 n} \left( \frac{\partial n}{\partial t} + \underset{\sim}{\nabla} \cdot n \underset{\sim}{v} \right) - \frac{D p}{D t}$$

$$= n \frac{D h}{D t} - h \left[ \underset{\sim}{\nabla} \cdot n \left( \underset{\sim}{w} - \underset{\sim}{v} \right) - \sum_j \frac{n_j}{m_j} \overline{\left( \delta m_j \right)^{coll}} \right] - \frac{D p}{D t}$$

where the mean number velocity

$$\underset{\sim}{w} := \frac{1}{n} \sum_j n_j \overline{\underset{\sim}{c}_j} \tag{4.12}$$

has been introduced.

The enthalpy balance is now

$$n\frac{Dh}{Dt} - \frac{Dp}{Dt} + \nabla\cdot\left[\sum_j \frac{\rho_j}{2}\overline{\underset{\sim}{C}_j C_j^2} - h\,n\left(\underset{\sim}{w}-\underset{\sim}{v}\right)\right]$$

$$+ n\left(\underset{\sim}{w}-\underset{\sim}{v}\right)\cdot\nabla h + \underset{\approx}{\pi} : \nabla\underset{\sim}{v} - \sum_j \dot{\underset{\sim}{c}}_{j0}\cdot\rho_j\overline{\underset{\sim}{C}_j} \qquad (4.13)$$

$$= \sum_j n_j\overline{\left(\frac{\delta}{\delta t}\frac{m_j}{2}C_j^2\right)^{coll}} - h\sum_j\frac{n_j}{m_j}\overline{\left(\delta m_j\right)^{coll}}$$

The first two terms can be written as

$$n\frac{Dh}{Dt} - \frac{Dp}{Dt} = \frac{Du}{Dt} - h\frac{Dn}{Dt} = n\left(\frac{D}{Dt}\frac{u}{n} + p\frac{D}{Dt}\frac{1}{n}\right).$$

With $1/n$ as the volume per particle this is the product of $n$ with the convective change of the heat per particle (first law of thermodynamics). The next term in the enthalpy balance (4.13) is the divergence of the total heat flux vector

$$\underset{\sim}{Q} := \sum_j \frac{\rho_j}{2}\overline{\underset{\sim}{C}_j C_j^2} - h\sum_j n_j\overline{\underset{\sim}{C}_j} = \sum_j n_j\left(\frac{m_j}{2}\overline{C_j^2\underset{\sim}{C}_j} - \frac{5kT}{2}\overline{\underset{\sim}{C}_j}\right), \qquad (4.14)$$

the other terms are dissipative terms. They are either bilinear in thermodynamic fluxes $n\left(\underset{\sim}{w}-\underset{\sim}{v}\right)$, $\underset{\approx}{\pi}$, $\rho_j\overline{\underset{\sim}{C}_j}$ and forces $\nabla h$, $\nabla\underset{\sim}{v}$, $\dot{\underset{\sim}{c}}_{j0}$ or caused directly by collisions.

According to the enthalpy balance (4.13) an adiabatic behaviour is obtained with the vanishing of the divergence of the total heat flux vector $\underset{\sim}{Q}$ (4.14). This is the sum of the partial heat flux vectors

$$\underset{\sim}{Q}_j := n_j\left(\frac{m_j}{2}\overline{C_j^2\underset{\sim}{C}_j} - \frac{5kT}{2}\overline{\underset{\sim}{C}_j}\right). \qquad (4.15)$$

The first terms are the flux vectors for the translational energies $m_j C_j^2/2$ and are therefore named partial translational energy flux vectors. The second terms are the partial enthalpy flux vectors $h\,n_j\,\overline{\underset{\sim}{C}_j}$.

For a one species gas the averaged peculiar velocity $\overline{\underset{\sim}{C}} = \overline{\underset{\sim}{c}} - \underset{\sim}{v}$ (2.14) (2.10) vanishes and the heat flux vector $\underset{\sim}{Q}$ equals the translational energy flux vector in this particular case. In a multi-species plasma, however, the averaged peculiar velocities $\underset{\sim}{C}_j = \underset{\sim}{c}_j - \underset{\sim}{v}$, multiplied with the charge densities $\rho_j^{el}$, are the partial conduction current densities $\underset{\sim}{j}_j = \rho_j^{el}\underset{\sim}{v}$ which cannot be neglected and therefore must be carried in the expressions (4.14) (4.15) for the heat flux vectors $\underset{\sim}{Q}$ and $\underset{\sim}{Q}_j$.

## 5. Balances for stresses and heat fluxes

For the calculation of transport coefficients for viscosity and heat conductivity balance equations for the partial stress tensors $\underset{\approx}{\pi}_j$ (3.13) and the partial heat flux vectors $\underset{\sim}{Q}_j$ (4.15) are needed. The corresponding quantities $\phi_j$ to be averaged are

$$m_j\underset{\sim}{C}_j\underset{\sim}{C}_j - kT\,\underset{\approx}{I} \qquad \frac{m_j}{2}C_j^2\underset{\sim}{C}_j - \frac{5kT}{2}\underset{\sim}{C}_j.$$

They depend on the (additional) macroscopic parameter $T(\underset{\sim}{r},t)$. The appropriate balance equations are special cases of (4.1) with

$$\overline{\underset{\sim}{C}_j\cdot\nabla\underset{\sim}{\phi}_j} + \frac{D\overline{\phi_j}}{Dt} = (\nabla T)\cdot\overline{\underset{\sim}{C}_j\frac{\partial\phi_j}{\partial T}} + \frac{DT}{Dt}\overline{\frac{\partial\phi_j}{\partial T}}. \qquad (5.1)$$

Putting $\phi_j = m_j \underset{\sim}{C}_j \underset{\sim}{C}_j - kT \underset{\approx}{I}$ we obtain the balances for the partial stresses

$$\frac{\partial}{\partial t} \underset{\approx}{\tau}_j + \underset{\sim}{\nabla} \cdot \left( \underset{\sim}{v} \underset{\approx}{\tau}_j + \underset{\approx}{Q}_j \right) + 2 \widehat{\underset{\sim}{\nabla} kT n_j \overline{\underset{\sim}{C}_j}} + n_j \overline{\underset{\sim}{C}_j} \cdot \underset{\sim}{\nabla} kT \underset{\approx}{I} + n_j \frac{D kT}{Dt} \underset{\approx}{I}$$

$$+ 2 \widehat{\underset{\approx}{\tau}_j \cdot \underset{\sim}{\nabla} \underset{\sim}{v}} + 2 n_j kT \widehat{\underset{\sim}{\nabla} \underset{\sim}{v}} + 2 \widehat{\left( \frac{D \underset{\sim}{v}}{Dt} - \underset{\sim}{\dot{c}}_{j0} \right) \rho_j \overline{\underset{\sim}{C}_j}} - 2 \widehat{\underset{\sim}{\Omega}_j \times \underset{\approx}{\tau}_j} \qquad (5.2)$$

$$= n_j \overline{\left( \frac{\delta m_j \underset{\sim}{C}_j \underset{\sim}{C}_j}{\delta t} \right)^{coll}} - kT \underset{\approx}{I} \frac{n_j}{m_j} \overline{\left( \frac{\delta m_j}{\delta t} \right)^{coll}}$$

with the symmetric third order partial heat flux tensors

$$\underset{\approx\approx}{Q}_j := n_j \left( m_j \overline{\underset{\sim}{C}_j \underset{\sim}{C}_j \underset{\sim}{C}_j} - 3 kT \widehat{\overline{\underset{\sim}{C}_j} \underset{\approx}{I}} \right) \qquad (5.3)$$

which are related to the partial heat flux vectors $\underset{\sim}{Q}_j$ (4.15) by

$$\underset{\sim}{Q}_j = \frac{1}{2} \, trace \, \underset{\approx\approx}{Q}_j \, . \qquad (5.4)$$

The definition for a symmetric third order tensor $\widehat{\underset{\approx\approx}{t}}$ is analogous to (4.7) for a symmetric second order tensor $\widehat{\underset{\approx}{t}}$, but now with 3! = 6 terms (at most) instead of two.

With $\phi_j = \frac{m_j}{2} C_j^2 \underset{\sim}{C}_j - \frac{5 kT}{2} \underset{\sim}{C}_j$ the following balance equations for the heat flux vectors $\underset{\sim}{Q}_j$ (4.15) are obtained:

$$\frac{\partial}{\partial t} \underset{\sim}{Q}_j + \underset{\sim}{\nabla} \cdot \left[ \underset{\sim}{v} \underset{\sim}{Q}_j + \underset{\approx}{R}_j + \frac{kT}{m_j} \left( \underset{\approx}{\tau}_j + \frac{1}{2} \underset{\approx}{I} \, trace \, \underset{\approx}{\tau}_j \right) \right]$$

$$+ \frac{5}{2} \underset{\approx}{p}_j \cdot \underset{\sim}{\nabla} \frac{kT}{m_j} + \frac{5}{2} \frac{D kT}{Dt} n_j \overline{\underset{\sim}{C}_j} + \left( \underset{\approx\approx}{Q}_j + 3 kT n_j \widehat{\overline{\underset{\sim}{C}_j} \underset{\approx}{I}} \right) : \widehat{\underset{\sim}{\nabla} \underset{\sim}{v}} + \underset{\sim}{Q}_j \cdot \underset{\sim}{\nabla} \underset{\sim}{v} \qquad (5.5)$$

$$+ \left( \frac{D \underset{\sim}{v}}{Dt} - \underset{\sim}{\dot{c}}_{j0} \right) \cdot \left( \underset{\approx}{\pi}_j + \frac{1}{2} \underset{\approx}{I} \, trace \, \underset{\approx}{\pi}_j \right) - \underset{\sim}{\Omega}_j \times \underset{\sim}{Q}_j$$

$$= n_j \overline{\left( \frac{\delta}{\delta t} \frac{m_j}{2} C_j^2 \underset{\sim}{C}_j \right)^{coll}} - \frac{5 n_j kT}{2 m_j} \overline{\left( \frac{\delta}{\delta t} m_j \underset{\sim}{C}_j \right)^{coll}}$$

The second order tensor $\underset{\approx}{R}_j$ is half of the trace of the symmetric fourth order tensor

$$\underset{\approx\approx}{R}_j := n_j \left[ m_j \overline{\underset{\sim}{C}_j \underset{\sim}{C}_j \underset{\sim}{C}_j \underset{\sim}{C}_j} - 6 kT \widehat{\overline{\underset{\sim}{C}_j \underset{\sim}{C}_j} \underset{\approx}{I}} + 3 \frac{(kT)^2}{m_j} \widehat{\underset{\approx}{I} \underset{\approx}{I}} \right] \qquad (5.6)$$

where again the definition for a symmetric fourth order tensor $\widehat{\underset{\approx\approx}{t}}$ is analogous to (4.7) for $\widehat{\underset{\approx}{t}}$, but with 4! = 24 terms (at most) instead of two.

## 6. Concluding remarks

The tensors $\rho_j \overline{\underset{\sim}{C}_j}$, $\underset{\approx}{\pi}_j$ (3.13), $\overset{Q}{\underset{\approx}{}}_j$ (4.15) and $\overset{R}{\underset{\approx}{}}_j$ (5.6) are averages of three-dimensional Hermite polynomials in the variables $\underset{\sim}{C}_j$. These Hermite polynomials are orthogonal with respect to the weight functions [4]

$$f_j^M(C_j^2, n_j, T) := n_j \left(\frac{m_j}{2\pi k T}\right)^{3/2} \exp - \frac{m_j C_j^2}{2kT} \qquad (6.1)$$

These are local Maxwellians. Plasmas in local thermodynamic equilibrium - see the description following eq. (3.11) - are completely described by these local Maxwellians. Consequently $\rho_j \overline{\underset{\sim}{C}_j}$, $\underset{\approx}{\pi}_j$, $\overset{Q}{\underset{\approx}{}}_j$, $\overset{R}{\underset{\approx}{}}_j$ and all the following averages of Hermite polynomials vanish in this limiting case.

For plasmas not in local thermodynamic equilibrium the distribution functions $f_j(\underset{\sim}{C}_j, \underset{\sim}{r}, t)$ can be expanded with the Hermite polynomials and the local Maxwellians (6.1) as weight functions. Then the tensors $\rho_j \overline{\underset{\sim}{C}_j}$, $\underset{\approx}{\pi}_j$, $\overset{Q}{\underset{\approx}{}}_j$, $\overset{R}{\underset{\approx}{}}_j$ etc. are proportional to the expansion coefficients. They decrease with increasing order (because of the orthogonality). To close the system of balance equations Grad [4] proposed to neglect all tensors of fourth and of higher order. To avoid a too abrupt truncation he further proposed to neglect also the tracefree parts

$$\overset{Q^0}{\underset{\approx}{}}_j := \overset{Q}{\underset{\approx}{}}_j - \frac{6}{5} \overline{\underset{\sim}{Q}_j \underset{\approx}{I}} \qquad (6.2)$$

of the heat flux tensors $\overset{Q}{\underset{\approx}{}}_j$ (4.15).

With these (or similar) neglects the distribution functions $f_j(\underset{\sim}{C}_j, \underset{\sim}{r}, t)$ are approximated. Then the collision integrals can be evaluated. This has been extensively done by Schunk and his collaborators. For a review see Schunk [5].

Another approach does not start mathematically with local Maxwellians (6.1) and therefore physically with local thermodynamic equilibrium, but with generalized local Maxwellians

$$f_j^G(\underset{\sim}{C}_j, n_j, \underset{\approx}{T}) := n_j \left(\frac{m_j}{2\pi k \det \underset{\approx}{T}}\right)^{3/2} \exp - \frac{m_j}{2k} \underset{\sim}{C}_j \cdot \underset{\approx}{T}^{-1} \cdot \underset{\sim}{C}_j \qquad (6.3)$$

with a symmetric temperature tensor $\underset{\approx}{T}(\underset{\sim}{r}, t)$ instead of the scalar temperature $T(\underset{\sim}{r}, t)$. For plasmas in strong magnetic fields $\underset{\sim}{B}$ the tensor $\underset{\approx}{T}$ is assumed to have the form

$$\underset{\approx}{T} = T_\parallel \underset{\sim}{\hat{B}}\underset{\sim}{\hat{B}} + T_\perp \left(\underset{\approx}{I} - \underset{\sim}{\hat{B}}\underset{\sim}{\hat{B}}\right). \qquad (6.4)$$

Then the generalized local Maxwellians (6.3) become the so-called bi-Maxwellians

$$f_j^B = n_j \left(\frac{m_j}{2\pi k T_\parallel}\right)^{1/2} \left(\frac{m_j}{2\pi k T_\perp}\right)^{2/2} \exp - \frac{m_j}{2k} \left(\frac{C_{j\parallel}^2}{T_\parallel} + \frac{C_{j\perp}^2}{T_\perp}\right). \qquad (6.5)$$

The generalized local Maxwellians (6.3) are weight functions of bi-orthogonal Hermite polynomials [6] in the variables $\underset{\sim}{C}_j$. Bi-Maxwellians (6.5) and the corresponding Hermite polynomials have been used to approximate distribution functions in plasmas pervaded by strong magnetic fields. For a review see Demars and Schunk [7].

References

[1] Landau, L.D., Lifschitz, E.M.: Elastizitätstheorie.
Berlin, Akademie-Verlag, 1965, eq. (2,3)

[2] Chapman, S., Cowling, T.G.: The mathematical theory of non-uniform gases.
Cambridge, University Press, 1970, eq. 8.1,3

[3] Suchy, K.: Calculation of collision integrals in the moment equations, in: Rarefied Gas Dynamics, Vol. I, (J.A. Laurman, ed.)
New York, Academic Press, 1963, pp. 181-208, eq. (61)

[4] Grad, H.: Note on N-dimensional Hermite polynomials; On the kinetic theory of rarefied gases.
Commun. pure appl. math. 2 (1949) 325-330; 331-407

[5] Schunk, R.W.: Mathematical structure of transport equations for multispecies flows.
Rev. Geophys. Space Phys. 15 (1977) 429-445

[6] Erdély, A. (ed.): Higher transcendental functions. Vol. II.
New York, McGraw Hill, 1953, sections 12.8 to 12.10

[7] Demars, H.G., Schunk, R.W.: Temperature anisotropies in the terrestrial ionosphere and plasmasphere.
Rev. Geophys. 25 (1987) 1659-1679

# Non-Linear Wave-Equations for Low-Frequency Acoustic Gravity Waves

R.H.M. Miesen, L.P.J. Kamp, F.W. Sluijter

Department of Physics, Eindhoven University of Technology,
5600 MB Eindhoven, The Netherlands.

**Abstract**

It is shown that low-frequency acoustic gravity waves propagating parallel to the earth's surface satisfy the Korteweg-De Vries equation or the Kadomtsev-Petviashvili equation and have a discrete spectrum of group velocities. The atmosphere is considered to be incompressible, homogeneous in composition and isothermal and the gravitational acceleration depends upon the height.

## 1. Introduction

In recent years there has been an increased interest in the properties of internal gravity waves in the earth's atmosphere (Ibrahim,Mostafa M., 1987; Stenflo, 1986, 1987; Weinstock, 1984; Yeh and Liu, 1981). This interest is partially motivated by the fact that these waves affect the refraction index of the atmosphere and consequently are relevant for various kinds of electromagnetic wave propagation (radio-astronomy, VLBI etc.)

Due to the exponential growth of the amplitude of internal waves in the earth's atmosphere with height, non-linear behaviour is likely to be expected. For this reason we have investigated the characteristic non-linear equations that govern low-frequency internal gravity waves in an incompressible isothermal atmosphere that is homogeneous in composition and in which the gravitational acceleration depends upon the height. Furthermore we use the hydrostatic approximation, which is valid in the low-frequency regime we have studied. We adapt this approximation in such a way that dispersion is not neglected entirely as is usually done in this approximation.

We have found that low-frequency internal gravity waves that propagate parallel to the surface of the earth should satisfy the well-known Korteweg-De Vries equation (in one space dimension) and the so-called Kadomtsev-Petviashvili equation (in two space dimensions). These equations apply to the horizontal and time variations of the disturbance. The vertical variation is described by Coulomb-wavefunctions.

Starting points are the equation of motion, the continuity equation and an equation of state. We have used spherical coordinates that have been modified in order to describe wave phenomena with wavelengths that are assumed to be small with respect to the circumference of the earth. So maximum wavelengths of several hundreds of kilometers are acceptable. The non-linear wave-equations were obtained in a rigorous way using reductive perturbation technique. A discrete spectrum of group velocities for these waves is found. Finally the modifications due to a compressible atmosphere are discussed.

## 2. Basic equations

Atmospheric wave motion is usually described by making use of the equations of continuity

$$\frac{\partial \rho}{\partial t} + \underline{v}\cdot\underline{\nabla}\rho = 0, \tag{1a}$$

$$\rho\left(\frac{\partial}{\partial t} + \underline{v}\cdot\underline{\nabla}\right)\underline{v} = -\underline{\nabla}p + \rho\underline{g}, \tag{1b}$$

$$\underline{\nabla}\cdot\underline{v} = 0, \tag{1c}$$

where $\rho$, p, $\underline{v}$, and g denote respectively mass density, pressure, fluid velocity and gravitational acceleration .

When the atmosphere is in equilibrium, (1b) reads (in spherical coordinates):

$$\frac{\partial p_o}{\partial r} + \rho_o g(r) = 0, \tag{2}$$

where $p_o$ and $\rho_o$ are the equilibrium pressure and mass density and r is the distance to the center of the earth. The gravitational acceleration is given by:

$$g(r) = g_o \frac{R^2}{r^2}, \tag{3}$$

where R is the radius of the earth and $g_o$ is the gravitational acceleration on the surface of the earth. From (2) and (3) together with the law for an ideal gas, i.e.

$$p_o = \rho_o \frac{KT}{m}, \tag{4}$$

where K is the Boltzmann-constant, T is the temperature of the atmosphere and m is the mean molecular mass, we obtain for the equilibrium mass density

$$\rho_o = \rho_{oo} \cdot \exp\left[\frac{R(R-r)}{r} H_o^{-1}\right], \tag{5}$$

where $\rho_{oo}$ is the mass density on the earthsurface, $H_o = (KT)/(mg_o)$ and m and T are assumed to be constant.

### 3. The hydrostatic approximation

Since we consider low-frequency acoustic gravity waves, i.e. acoustic gravity waves with frequencies considerably smaller than the Brunt-Väisälä frequency $\omega_b$, where

$$\omega_b^2 = \frac{g}{H}, \tag{6}$$

we can use the so-called hydrostatic approximation (Gill, 1982).

In this approximation the pressure at any point is approximated by the weight of a unit cross-section of air above that point. In formula

$$\frac{\partial p}{\partial r} = -\rho g. \tag{7}$$

This means that in the hydrostatic approximation the vertical acceleration in (1b) is neglected. This will turn out to be essential in the derivation of a non-linear wave-equation for acoustic gravity waves.

### 4. The Korteweg-De Vries equation

Equations (1) read in spherical coordinates with $\partial/\partial\varphi$ chosen to be zero

$$\frac{\partial \rho}{\partial t} + \frac{u}{r}\frac{\partial \rho}{\partial \Theta} + w\frac{\partial \rho}{\partial r} = 0, \tag{8a}$$

$$\rho\left(\frac{\partial}{\partial t} + \frac{u}{r}\frac{\partial}{\partial \Theta} + w\frac{\partial}{\partial r}\right)u + \frac{1}{r}\frac{\partial p}{\partial \Theta} = 0, \tag{8b}$$

$$\rho\left(\frac{\partial}{\partial t} + \frac{u}{r}\frac{\partial}{\partial \Theta} + w\frac{\partial}{\partial r}\right)w + \frac{\partial p}{\partial r} + g\rho = 0, \tag{8c}$$

$$\frac{1}{r^2}\frac{\partial}{\partial r}(r^2 w) + \frac{1}{r\sin(\Theta)}\frac{\partial}{\partial \Theta}[u\sin(\Theta)] = 0, \tag{8d}$$

where u is the azimuthal and w is the radial velocity.

The kind of waves we are considering have dimensions that are much smaller than the circumference of the earth. Therefore

$$\sin(\theta)\,\frac{\partial u}{\partial \theta} \gg u\,\frac{\partial}{\partial \theta}\,[\sin(\theta)], \tag{9}$$

and (8d) yields

$$\frac{\partial w}{\partial r} + \frac{2}{r}w + \frac{1}{r}\frac{\partial u}{\partial \theta} = 0. \tag{8d'}$$

We choose r and r·sin(θ) as new independent variables, so that

$$\left.\frac{\partial}{\partial r}\right|_{\theta} \rightarrow \left.\frac{\partial}{\partial r}\right|_{x} + \frac{x}{r}\left.\frac{\partial}{\partial x}\right|_{r}, \tag{10a}$$

$$\left.\frac{\partial}{\partial \theta}\right|_{r} \rightarrow r\left(1 - \frac{x^2}{r^2}\right)^{1/2}\left.\frac{\partial}{\partial x}\right|_{r}. \tag{10b}$$

Because in our problem

$$x \ll R \leq r \tag{11}$$

we can approximate (10) by

$$\left.\frac{\partial}{\partial r}\right|_{\theta} \rightarrow \left.\frac{\partial}{\partial r}\right|_{x}, \tag{12a}$$

$$\left.\frac{\partial}{\partial \theta}\right|_{r} \rightarrow r\left.\frac{\partial}{\partial x}\right|_{r}. \tag{12b}$$

When we use (12), (8) becomes

$$\rho_t + \rho u_x + \rho w_r = 0, \tag{13a}$$

$$\rho(u_t + uu_x + wu_r) + p_x = 0, \tag{13b}$$

$$\rho(w_t + uw_x + ww_r) + p_r + g\rho = 0, \tag{13c}$$

$$w_r + \frac{2}{r}w + u_x = 0. \tag{13d}$$

where subscripts denote differentiations and u is now the x-component of the velocity.

To derive a non-linear equation for acoustic gravity waves we will use a technique that is commonly used in studying long, non-linear, unidirectional waves (Dodd, 1982). This technique is called reductive perturbation technique and combines expansion of the dependent variables in terms of a perturbation parameter with, at the same time, rescaling of the independent variables.

Using this method we choose following new independent variables

$$\xi = x - at, \tag{14a}$$

$$\tau = \epsilon t, \tag{14b}$$

where a is a certain speed and $\epsilon$ is a small parameter ($\epsilon \ll 1$).

The dependent variables are expanded as follows

$$\frac{\rho}{\rho_o} = 1 + \epsilon\,\rho_1 + \epsilon^2\rho_2 + \ldots \tag{15a}$$

$$\frac{p}{g_o H_o \rho_o} = 1 + \epsilon\, p_1 + \epsilon^2 p_2 + \ldots, \tag{15b}$$

$$\underline{v} = \epsilon \underline{v}_1 + \epsilon^2 \underline{v}_2 + \ldots, \tag{15c}$$

where we have assusmed that $\underline{v}_o = \underline{0}$, which means that wind is not included. Equations (14) and (15) are substituted in (13). We adapt the hydrostatic approximation in a way that terms that correspond to vertical acceleration are assumed to be of an order in $\epsilon$ smaller than other terms in (1b). We now find the following equation for terms of the first order in $\epsilon$

$$a\rho_{1\xi} + \frac{R^2}{r^2} H_o^{-1} w_1 = 0, \tag{16a}$$

$$-a u_{1\xi} + g_o H_o p_{1\xi} = 0, \tag{16b}$$

$$g_o H_o p_{1r} - \frac{R^2}{r^2} g_o p_1 + \frac{R^2}{r^2} g_o \rho_1 = 0. \tag{16c}$$

$$w_{1r} + \frac{2}{r} w_1 + u_{1\xi} = 0. \tag{16d}$$

Where we have used the fact that

$$\rho_o^{-1} \frac{\partial \rho_o}{\partial r} = H_o^{-1} \frac{R^2}{r^2}. \tag{17}$$

Equations (16) yield the following differential equation for $w_1$:

$$w_{1rr} + \left[\frac{2}{r} - \frac{R^2}{r^2} H_o^{-1}\right] w_{1r} + \left[\omega_{bo}^2 a^{-2} \frac{R^4}{r^4} - \frac{2R^2}{r^3} H_o^{-1} - \frac{2}{r^2}\right] w_1 = 0. \tag{18}$$

where

$$\omega_{bo}^2 = \frac{g_o}{H_o}.$$

To solve this equation, we define functions $\tilde{w}_1(\xi,\tau)$ and $h(r)$ as

$$w_1(\xi,r,\tau) = h(r)\tilde{w}_1(\xi,\tau). \tag{19}$$

The function $h(r)$ has to satisfy the same differential equation as $w_1(\xi,r,\tau)$ (i.e. (18)). The solution for $h(r)$ that remains finite for r to infinity is

$$h(r) = \alpha \cdot \exp\left[-\frac{R^2}{2H_o r}\right] \cdot F_1(\eta, \tfrac{\nu}{r}), \tag{20}$$

where $F_1(\eta,\frac{\nu}{r})$ is a Coulomb-wavefunction (Abramowitz and Stegun, 1972), $\alpha$ is an integration constant and

$$\nu = R^2\left[\frac{\omega_{bo}^2}{a^2} - \frac{1}{4H_o^2}\right]^{1/2}. \tag{21}$$

$$\eta = R^2 (H\, \nu)^{-1}. \tag{22}$$

The boundary condition

$$w_1(\xi, r=R, \tau) = 0, \tag{23}$$

due to the infinitely stiff earth surface implies that

$$F_1(\eta, \frac{\upsilon}{R}) = 0. \tag{24}$$

This equation determines discrete values for the velocity a that are possible. Notice that both $\eta$ and $\upsilon$ contain a. Therefore (24) is quite difficult to solve.

Another point of interest is that in the usual flat-earth linear theory there is but one mode that does meet the condition of zero vertical velocity at the surface of the earth, under the isothermal, windfree conditions assumed here, and that is the Lamb mode with horizontal speed of propagation equal to the speed of sound. An important difference between the modes determined by (24) that satisfy this condition and the Lamb mode is that for the Lamb mode the vertical velocity is zero for the whole atmosphere and that is certainly not true for the modes determined by (24).

We now proceed with the equations for terms of the second order in $\epsilon$ to determine $\tilde{w}_1(\xi, \tau)$. These equations are

$$\rho_{1\tau} - a\rho_{2\xi} - \frac{R^2}{r^2}H_o^{-1}w_2 - \frac{R^2}{r^2}H_o^{-1}\rho_1 w_1 + u_1\rho_{1\xi} + w_1\rho_{1r} = 0, \tag{25a}$$

$$u_{1\tau} - au_{2\xi} + u_1u_{1\xi} + w_1u_{1r} + g_oH_o\rho_{2\xi} = 0, \tag{25b}$$

$$-aw_{1\xi} + g_oH_o\rho_{2r} - \frac{R^2}{r^2}g_o\rho_2 + \frac{R^2}{r^2}g_o\rho_2 = 0, \tag{25c}$$

$$w_{2r} + \frac{2}{r}w_2 + u_{2\xi} = 0. \tag{25d}$$

Solving $\rho_{2\xi}$ from (25a), $p_{2\xi}$ from (25b) and substituting these two in (25c) gives, together with (25d), an equation for $u_1, w_1, \rho_1$ and $w_2$. From this equation $\rho_1$ and $u_1$ can be eliminated by using (16a) and (16d). We then obtain the following equation for $w_2$ :

$$w_{2rr} + \left\{\frac{2}{r} - \frac{R^2}{r^2}H_o^{-1}\right\}w_{2r} + \left\{\omega_{bo}^2a^{-2}\frac{R^4}{r^4} - \frac{2R^2}{r^3}H_o^{-1} - \frac{2}{r^2}\right\}w_2 = -2\omega_{bo}^2a^{-3}\frac{R^4}{r^4}h\tilde{\varphi}_{1\tau} + a^{-1}\tilde{\varphi}_{1\xi} +$$

$$+ \left\{h_{rrr} - h_{rr}\left[H_o^{-1}\frac{R^2}{r^2} + \frac{2}{r}\right] + h_r\left[\omega_{bo}^2a^{-2}\frac{R^4}{r^4} - \frac{12}{r^2}\right] + h\left[\omega_{bo}^2a^{-2}\frac{R^6}{r^6}H_o^{-1} + \omega_{bo}^2a^{-2}\frac{2R^4}{r^5} + \right.\right.$$

$$\left.\left. + \frac{6R^2}{r^4}H_o^{-1} + \frac{12}{r^3}\right]\right\} - h\tilde{\varphi}_{1\xi\xi\xi} \tag{26}$$

where $\tilde{\varphi}_{1\xi} = \tilde{w}_1$.

When we multiply (26) with $hr^{-2}\exp[R^2/H_or)]$ and integrate from $r = R$ to infinity, the left hand side of this equation is zero. So the right hand side of the equation has to be zero too. Thus we obtain the final result

$$I_1\tilde{\varphi}_{1\tau} + I_2\tilde{\varphi}_1\tilde{\varphi}_{1\xi} + I_3\tilde{\varphi}_{1\xi\xi\xi} = 0, \tag{27}$$

where

$$I_1 = \int_R^\infty 2h^2\omega_{bo}^2 a^{-3} \frac{R^4}{r^6} \exp\left(\frac{R^2}{H_o r}\right)dr,$$

$$I_2 = \int_R^\infty a^{-1}h^2 r^{-2}\left\{-h_{rrr} + h_{rr}\left[\frac{R^2}{r^2}H_o^{-1} + \frac{2}{r}\right] + h_r\left[-\omega_{bo}^2 a^{-2}\frac{R^4}{r^4} + \frac{12}{r^2}\right] - h\left[\omega_{bo}^2 a^{-2}\frac{R^6}{r^6}H_o^{-1} + \right.\right.$$
$$\left.\left. + \omega_{bo}^2 a^{-2}\frac{2R^4}{r^5} + \frac{5R^2}{r^4}H_o^{-1} + \frac{12}{r^3}\right]\right\}\cdot\exp\left[\frac{R^2}{H_o r}\right]dr,$$

$$I_3 = \int_R^\infty h^2 r^{-2}\exp\left[\frac{R^2}{H_o r}\right]dr.$$

Equation (27) is the well-known Korteweg-de Vries equation (Korteweg and De Vries, 1895). In our model we did not take any dissipative effects into account and consequently the waves extend ad infinitum, which is also the case in ordinary linear theory. In order to obtain a more realistic value of the integrals in (27), one can integrate from R to the height where these dissipative effects become dominating.

**5. The Kadomtsev-Petviashvili equation**

In the preceding we have assumed that differentiations with respect to $\varphi$ are zero. That means that the acoustic gravity wave is uniform in that direction. If we abandon this assumption we find another non-linear wave-equation.

We now assume the velocity in the y-direction, i.e. v, to be considerably smaller than the velocities in the x- and r-direction, i.e. u and w. With this assumption the equations for the terms of the first order in $\epsilon$ are still given by (16). From the equations for the terms of the second order in $\epsilon$, only (25d) changes into

$$w_{2r} + \frac{2}{r}w_2 + u_{2\xi} + v_{1y} = 0, \tag{28}$$

where $v_{1y}$ follows from the y-component of the momentum equation, i.e.

$$av_{1y} = g_o H_o p_{1y}. \tag{29}$$

Equations (16), (25), (28) and (29) yield, in the same way as in the preceeding section the so-called Kadomtsev-Petviashvili equation, (Kadomtsev and Petviashvili, 1970)

$$(I_1\tilde{\varphi}_{1\tau} + I_2\tilde{\varphi}_1\tilde{\varphi}_{1\xi} + I_3\tilde{\varphi}_{1\xi\xi\xi})_\xi + \tfrac{1}{2}aI_1\tilde{\varphi}_{1yy} = 0, \tag{30}$$

where $I_1$, $I_2$ and $I_3$ are still given by (27).

**6. A compressible atmosphere**

Atmospheric wave motions in a compressible atmosphere are usually described by using the equations of momentum and continuity, together with an equation of state (e.g. Yeh and Liu, 1972)

$$\frac{\partial\rho}{\partial t} + \underline{\nabla}\cdot(\rho\underline{v}) = 0, \tag{31a}$$

$$\rho\left(\frac{\partial}{\partial t} + \underline{v}\cdot\underline{\nabla}\right)\underline{v} = -\nabla p + \rho\underline{g}, \tag{31b}$$

$$\rho\left(\frac{\partial}{\partial t} + \underline{v}\cdot\underline{\nabla}\right)(p\rho^{-\gamma}) = 0, \tag{31c}$$

where $\gamma$ is the ratio of specific heats.

In the same way we obtained (16) we now find

$$-a\rho_{1\xi} - \frac{R^2}{r^2}H_o^{-1}w_1 + w_{1r} + \frac{2}{r}w_1 + u_{1\xi} = 0, \tag{32a}$$

$$-au_{1\xi} + g_oH_op_{1\xi} = 0, \tag{32b}$$

$$g_oH_op_{1r} - \frac{R^2}{r^2}g_op_1 + \frac{R^2}{r^2}g_o\rho_1 = 0, \tag{32c}$$

$$-a(g_oH_op_{1\xi} - c^2\rho_{1\xi}) + \frac{R^2}{r^2}g_o(\gamma-1)w_1 = 0, \tag{32d}$$

The differential equation (18) becomes

$$w_{1rr} + \left[\frac{2}{r} - \frac{R^2}{r^2}H_o^{-1}\right]w_{1r} + \left[\omega_{go}^2a^{-2}\frac{R^4}{r^4} - \frac{2(1-2/\gamma)R^2}{r^3}H_o^{-1} - \frac{2}{r^2}\right]w_1 = 0, \tag{33}$$

where $\omega_{go}^2 = g_o^2(\gamma-1)/c^2$.

Notice that when $\gamma$ goes to infinity, i.e. the atmosphere becomes incompressible and the speed of sound goes to infinity, that (33) reduces to (18) and $\omega_{go}$ approaches $\omega_{bo}$.

The solutions of (33) are of the same type as those of (18). Instead of (25) we now find

$$\rho_{1\tau} - a\rho_{2\xi} - \frac{R^2}{r^2}H_o^{-1}w_2 - \frac{R^2}{r^2}H_o^{-1}\rho_1w_1 + u_1\rho_{1\xi} + w_1\rho_{1r} + w_{2r} + \frac{2}{r}w_2 + u_{2\xi} + \rho_1w_{1r} + \frac{2}{r}\rho_1w_1 + \\ + \rho_1u_{1\xi} = 0, \tag{34a}$$

$$u_{1\tau} - au_{2\xi} + u_1u_{1\xi} + w_1u_{1r} + g_oH_op_{2\xi} = 0, \tag{34b}$$

$$-aw_{1\xi} + g_oH_op_{2r} - \frac{R^2}{r^2}g_op_2 + \frac{R^2}{r^2}g_o\rho_2 = 0, \tag{34c}$$

$$(g_oH_op_{1\tau} - c^2\rho_{1\tau}) - a\{g_oH_op_{2\xi} - c^2\rho_{2\xi} + c^2\rho_1\rho_{1\xi} - \gamma g_oH_o(p_1\rho_1)_\xi\} + u_1(g_oH_op_{1\xi} - c^2\rho_{1\xi}) \\ + w_1(g_oH_op_{1r} - c^2\rho_{1r}) + (\gamma-1)w_1(g_oH_op_1 - c^2\rho_1)\frac{R^2}{r^2}H_o^{-1} + \frac{R^2}{r^2}g_o(\gamma-1)w_2 = 0. \tag{34d}$$

Proceeding in the same way as we did for the incompressible atmosphere we find from (34) an equation that differs from (27) only in the value of $I_2$. The integral $I_2$ then becomes a significantly more complicated, but nevertheless calculable expression. Within the assumptions that have been made, the only physical consequence is that, beside of the correction of the Brunt-Väisälä frequency, the strength of the non-linearity in the Korteweg-De Vries equation changes. This however can be accounted for by a rescaling of the dependent variable $\tilde{\varphi}$ in (27).

## 7. Remarks

The constant $\alpha$ in (20) can't be chosen arbitrary because the reductive perturbation technique is only valid for weakly non-linear problems. When $\alpha$ is too large the solutions of our problem would not satisfy this starting condition that the ratio of the perturbation and the equilibrium quantities has to be of order $\epsilon$.

From (21) we see that when $|a| > 2H_o\omega_{bo}$, $\nu$ is imaginary. However, for an imaginary argument, the Coulomb-wavefunctions do not possess any zeros. In that case it is not possible to satisfy the boundary condition (23). Therefore $|a|$ has to be smaller than $2H_o\omega_{bo}$. It is interesting to notice that this is also the maximum group velocity for

gravity waves as it follows from the dispersion relation for the linearised eqs. (1). This maximum group velocity is in turn smaller than the sound speed $\gamma KT/m$.

A solution of (30) is (Dodd, 1982)

$$\tilde{\varphi}_1(\xi,\tau) = \frac{AI_2}{12I_1}\left[\frac{I_1}{I_3}\right]^{1/3}\cosh^{-1}\left\{(A^{1/2}/2)\left[\left[\frac{I_1}{I_3}\right]^{1/3}\xi - A\tau + \varphi\right]\right\}, \tag{35}$$

where A and $\varphi$ are arbitrary real constants.

Therefore we find for $w_1$

$$w_1(\xi,r,\tau) = \frac{v^{3/2}I_2}{24I_1}\left[\frac{I_1}{I_3}\right]^{2/3}\exp\left(-\frac{1}{2H_o r}\right)\cdot F\left(\eta,\frac{v}{r}\right)\cdot$$
$$\cdot\cosh^{-2}\left\{(A^{1/2}/2)\left[\left[\frac{I_1}{I_3}\right]^{1/3}\xi - A\tau + \varphi\right]\right\}\sinh\left\{(A^{1/2}/2)\left[\left[\frac{I_1}{I_3}\right]^{1/3}\xi - A\tau + \varphi\right]\right\}, \tag{36}$$

where we have used (19).

**References**


Abramowitz, M. and Stegun, A., 1972: Handbook of mathematical functions. Dover publications, New York, 537-554.

Dodd, R.K., 1982: Solitons and Nonlinear Wave Equations, Academic Press, 4.

Gill, A.E., 1982: Atmosphere-Ocean Dynamics. Academic Press.

Ibrahim, Mostafa M., 1987: Effect of Nonlinearity on Atmospheric Gravity Waves, J. Atmos. Sci. 44, 706-720.

Kadomtsev, B.B. and Petviashvili, V.I., 1970: On the stability of solitary waves in weakly dispersing media. Sov. Phys. Doklady 15, 539-541.

Korteweg, D.J. and De Vries, G, 1895: On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves. Phil.Mag. 39, 422-443.

Stenflo, L., 1986: Nonlinear Equations for Acoustic Gravity Waves, Physica Scripta 33, 156-158.

Stenflo, L., 1987: A Nonlinear Equation for Low-Frequency Acoustic Gravity Waves. Physica Scripta 35, 488-489.

Weinstock, J., 1984: Gravity Wave Saturation and Eddy Diffusion in the Middle Atmosphere. J. Atmos. Terr. Phys. 46, 1069-1082.

Yeh, K.C. and Liu, C.H., 1981, J. Geophys. Res. 86, 9722.

Yeh, K.C. and Liu, C.H., 1972: Theory of Ionospheric Waves. Academic Pres, New York, 403-413.

## DISCUSSION

**G.Rostoker**

Is there a physical justification for the assumption that $\delta/\epsilon = 0(1)$, or was it a necessary assumption in order for the equations to be analytically soluble?

**Author's Reply**

To both questions the answer is yes. We need this assumption to solve the equations, but there is also a physical justification for this assumption.

In the field of non-linear wave propagation, both experimentally and theoretically, it is well known that when we have dispersion and also non-linearity a balance will be created between these two things for long waves. Long, in this context, means waves with horizontal wavelengths larger than the scale height (H) of the atmosphere. Balancing these two effects is exactly the consequence of assuming $\delta/\epsilon$ to be of order one.

We can physically justify this assumption also by saying that we are studying waves for which $\delta/\epsilon$, i.e. $(2H\omega_a/c)^2\epsilon^{-1} = 0(\omega_0^2/\omega_b^2\epsilon)$ is of order one; so waves for which the characteristic frequency is of the order $\epsilon^{1/2}\omega_b$. This frequency range is indeed the one for the long internal gravity waves we are studying.

K.C.Yeh

(a) In one of your transparencies you show possibilities of three nonlinear wave processes. Of course, since we are dealing with the same set of hydrodynamic equations, at a given environment, there can be only one physical phenomenon. Can you describe the conditions under which your condition for propagation of a solitary wave will manifest?

(b) In response to my question (a), you seem to indicate that the excitation of solitary waves depends only on the source conditions. Is this true?

**Author's Reply**

(a) In general, it is quite difficult to say when acoustic gravity waves will be excited. We expect these waves to manifest when we have some source, which excites internal gravity waves with a frequency that is considerably smaller than the Brunt-Väisälä frequency (though not too small because then the rotation of the earth becomes important and we have neglected this rotation). These waves have been observed experimentally, generated by Tsanumi waves or ionospheric substorms. They also have been observed, e.g., above Australia, but the source was not detected.

(b) When there is a source generating internal gravity waves with the considered frequencies in a more or less horizontal direction, the only other condition for the formation of the solitary waves considered by us is the condition that we can consider the medium to be infinite in the direction parallel to the surface of the earth.

# GLOBAL IONOSPHERIC DYNAMICS

K. Rawer
Albert-Ludwigs-Universität
D-7800 Freiburg
Fed. Rep. of Germany

## SUMMARY

Plasma being a minor constituent in planetary atmospheres its motion is narrowly linked with that of the neutrals as long as collisional coupling between neutral and charged particles is strong, i.e. in the lower layers. At greater heights, however, both constituents are more and more decoupled and may move separately whenever forces are present which do only act upon charged particles, i.e. electric or magnetic fields.

The most important difference between known planetary ionospheres is the absence or presence of a general magnetic field. While in the first condition the plasma moves freely, a magnetic field cannot provoke plasma motions of its own but restricts the motions provoked by other forces to one in either the direction parallel to the magnetic field - so for mechanical forces - or perpendicular to it - for electric forces. So, the motion patterns depend largely on the prevailing forces.

In this respect, largely different conditions are found in the terrestrial ionosphere according to magnetic latitude. While at mid-latitudes mechanical forces mainly provoked by neutral winds are most important, electric forces cannot be neglected at low and high magnetic latitudes. Apart from such large scale motions, one observes also many small scale motions of more or less wave-like character, in particular internal gravity waves due to different causes. It is not always easy in local measurements to separate these localized motions from the large scale patterns.

There exist different methods for observing plasma motions in the terrestrial ionosphere. Unfortunately, all of these can only be applied with some caution. Also, the spatial and temporal density of such local observations is unsatisfactory. Satellites could not yet produce large-scale surveys either, but observations relying on aircraft have allowed regional studies to be carried our at least.

## 1. INTRODUCTION

Everywhere in the terrestrial ionosphere the ionized components, electrons and ions, are minor constituents. Provided coupling with the neutral particle population is strong, the ionized populations are carried with the neutral wind, coupling being due to frequent collisions. In that situation, "plasma dynamics" is identical with atmosphere dynamics. So, at lower altitudes we do not have separate motions of the ionized components. These latter, therefore, can be used as "tracers" of the neutral motion. This is, for example, done when determining the motion of the ionized traces of meteors with radio methods. The height range where collisional coupling is strong certainly covers the mesosphere and reaches up to about the 100 km level.

In the upper height range, where collisions are less frequent, the motion of the ions can be different from the neutral motion. The free electrons, however, are not so free that they could move independently. Since their mass is small compared with that of the ions, these latter easily pull the electrons with them. Thus, ionospheric dynamics essentially means motion of ions.

Of course, above the 100 km level, coupling with neutral motions is of great importance too. Momentum transfer from neutral particles is still a most important force. However, in that upper height range electric and magnetic forces also influence the ion motion such that the effective motions might differ from the direction of the average momentum transfer, i.e. from neutral wind or from neutral oscillatory motions as occur, for example, in gravity waves.

Let us first consider an atmosphere in which electric and magnetic forces play a minor role. Venus is a good example. Energy input by solar wave radiation occurs on the daylight side where a pressure 'High' forms near the noon meridian. Pressure gradient forces are directed away from the "noon bulge", i.e. they must point towards the "terminator". So, the neutral motion is mainly zonal, neutrals and plasma move together. Therefore, (at lower latitudes at least) we have an eastward wind in the morning and a westward one in the afternoon. The plasma is carried with these winds so that the nightside of the planetary ionosphere receives an influx of ions and electrons from both sides. Both fluxes collide at the midnight meridian where the ordered motions are stopped, their kinetic energy reappearing as heat.

Plasma measurements in the Venus ionosphere were carried out aboard NASA's Venus Orbiter with a "retarding potential analyzer" (Knudsen et al., 1979, 1980, 1981). On the day side ion velocities (towards the terminator) as large as 10 km/s were measured. On the night side the Venus ionosphere has much lower density than by day, but the temperature vs. height profile remains the same as on the day side - except for the midright meridian, where the highest ion temperatures were found (Figure 1). These findings are in agreement with the explanation given above.

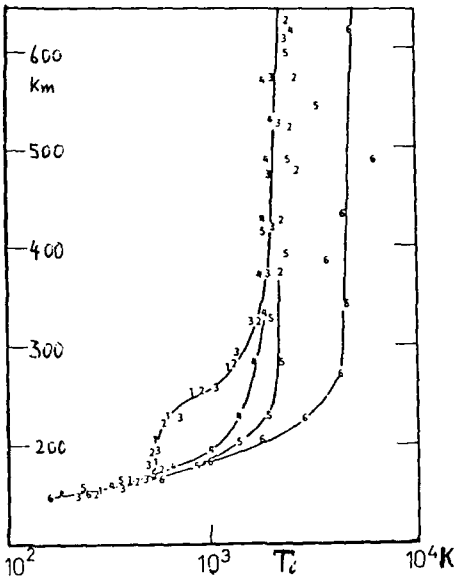



Figure 1. Mean ion temperature profiles in the Venus ionosphere, ordered after the solar zenith angle: numbers 1 .. 3 day, 4 .. 6 night; 6 is near the anti-solar meridian. (Courtesy: W.C. Knudsen and K. Spenner).

Another, typical feature of the Venus ionosphere is the fact that it is scarcely protected from the solar wind. This latter reaching down to a few 100 km above surface, it acts as an important heat source at the top of the Venusian ionosphere.

Planet Earth, unfortunately, has quite different conditions than Venus has:

1- Earth rotates quite fast (while Venus has a rotation period of 243 d). Therefore, Coriolis-forces are quite important (in the usual, "Earth-bound", frame).

2- Earth has an important magnetic field, $\underline{B}$, which has three main effects: (i) It is a shield against the solar wind; (ii) it has a strong effect on plasma motions of all kinds (except the "field-aligned" ones); (iii) in the presence of an electric field, $\underline{E}$, it provokes a crossfield drift, $\underline{E} \times \underline{B}/B^2$ (of same direction for charges of both signs).

Since the magnetic field exerts a "guiding" action on plasma motions, mechanical forces can only shift the plasma along the field lines, not across. This holds as long as the free path of the ions is much greater than the length of the gyro-circle, i.e. when the ion gyro frequency is much higher than their collision frequency.

Coupling with neutral atmosphere wind is by collisions of neutral molecules. In the average of many collisions these, finally, shift the plasma as whole, up or down the fieldline. This fact has a striking effect on the plasma motions provoked by pressure gradients: zonal motions are suppressed while meridional ones are allowed. In the magnetic meridian plane plasma transport only can occur along the fieldlines. Therefore, in spite of the fact that they are situated on different hemispheres, the ionospheric conditions at magnetically conjugate points can be stronger coupled than at neighbouring locations.

Such reasonning applies to all non-electric forces. The motions, such forces intend to provoke, are turned around by the magnetic field that exerts a Lorentz-force onto any moving charge. The ions do, therefore, move on helical orbits around the fieldline, they cannot come away from it except when collisions are frequent enough. Quite different is the effect of an external electric field. The field component parallel to the magnetic field provokes an acceleration in that direction, proportional to $E_{\parallel}$. On the other hand, we have, of course, the $\underline{E} \times \underline{B}/B^2$ drift shifting ions and electrons across the magnetic field lines. In particular, electric fields in the meridional plane provoke a zonal motion of the plasma. And vice versa: whenever a zonal plasma motion is observed, we conclude that there must be an electric field present. This is the only exception from the otherwise "field-aligned" plasma motion.

The field-aligned motion of ions in the terrestrial magnetic field produces a characteristic fibrous structure of the ionospheric topside plasma. Since diffusive exchange is hindered in cross-field directions, an inhomogeneous horizontal density distribution, when existing at the base level, must reproduce itself at greater heights. This "spaghetti" structure is the reason why whistlers often propagate on different neighbouring "orbits" so that multiple traces appear in sonagrams. Since the "spaghettis" are well conducting, electrically the medium behaves like a bundle of parallel wires, insulated against each other, such that crosswise conduction is quite poor.

## 2. REACTION TO EXTERNAL MECHANICAL FORCES

### 2.1 Gravity, Sedimentation, Diffusion

There were no atmospheres without gravity. It is the reason of sedimentation which, together with thermal diffusion, determines the atmospheric structure. The motion of an individual neutral particle between subsequent collisions is free. The mean free path increases with increasing altitude from about 1 m around the 100 km level to 1 km around 300 km and more than 100 km in the exosphere. There, the air molecules that have a positive vertical velocity component perform extended parabolic orbits, coming down with the same kinetic energy but inverted vertical component.

Not so for ions. Since these are bound to their magnetic fieldline, they move up and down that line, or, if their kinetic energy is high enough, they reach the peak of the fieldline and come down in the other hemisphere. In order to do so - according to the L-value - a proton needs a kinetic energy between .1 and .5 eV. (An $O^+$ ion needs 16 times more; that energy cannot be acquired without special acceleration, e.g. by electrodynamic forces).

As a transport process, diffusion - in a gravity field - tends to achieve sedimentation, individually for each atmospheric component. The relevant scale height depends on the component mass and temperature. As for plasma components, however, electrons are not independent from ions, because strong electrostatic forces build-up between charges of both signs. On the other hand - except for lower heights where collisional coupling is important - the different ion species find their distribution independently from each other. The mechanism is explained by Figure 2a.

There is one important difference against the neutrals: ionized components are not a stable population because they disappear by recombination. This latter is almost negligible at greater heights but quite frequent in the lower thermosphere. So, this lower height range is a plasma sink; in the net average, there is a downward plasma flux with plasma annihilation at the base of the ionosphere (in spite of the very important daytime ion production at that level). Thus, different from the stable neutral components for which the component density increases monotoneously downwards, the ion populations exhibit a drastic decrease at the base of the ionosphere.

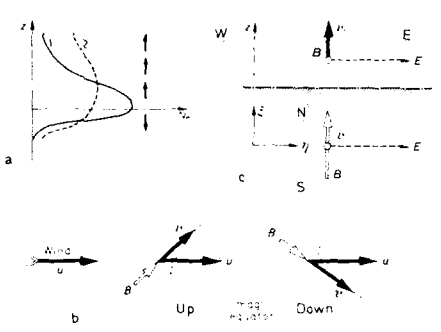



Figure 2. Transport mechanisms in the F-region by (a) diffusion: shape 1 toward shape 2 /velocities appear as arrows/; (b) momentum transfer from a neutral (horizontal) wind, $\underline{u}$: ions move with $\underline{v}_i$; (c) electric field, $\underline{E}$, causing a cross-field drift $\underline{E} \times \underline{B}/B^2$. ($\underline{v}_i$ is therefore vertical at the magnetic equator, where $\underline{B}$ is horizontal) (Rawer, 1984).

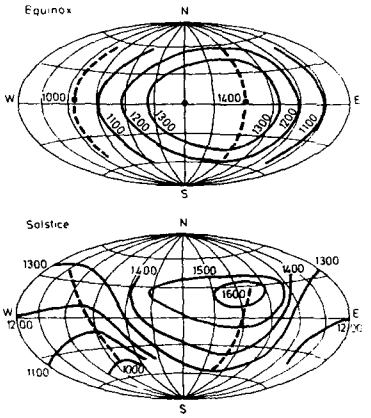



Figure 3. Crude maps of exospheric temperature inferred from incoherent scatter data at 12°S and 45°N. (Waldteufel et al., 1969).

## 2.2 Neutral winds

Momentum transfer from the neutral components is the most important mechanism of plasma transport, particularly at middle heights and latitudes. Since each ion is - more or less - fixed to its magnetic fieldline, the only large-scale induced motion is along that fieldline. Whether it is "up" or "down" depends on the configuration - as shown by Figure 2b. This is a mechanism explaining a large part of the observed variations of ionospheric peak parameters. Note that shifting down means strengthening, shifting up weakening of ion recombination (Kohl and King, 1965; Rishbeth et al., 1965).

In this context it is important to know the neutral winds. Earlier computations used exospheric temperature maps as given in CIRA-72. Inverting the above considerations, temperature maps were also derived from ion temperatures obtained with the incoherent scatter technique (Figure 3). In all these maps, a strong diurnal tidal component should be the most important feature. The observations, however, show a more involved pattern, in which the semi-diurnal components are quite strong (Figures 4 to 6). Comparable tidal effects were also detected in the composition of the neutral atmosphere (Hedin et al., 1980).

In the lower thermosphere, other methods of observation can be used: tristatic observations of sound propagation from rockets, Radar following of artificial tracers ("chaffs") and of natural ones (plasma irregularities, meteor trails, airglow features). In the recent time, the meteor Radar-technique is applied at many mid-latitude stations. These observations revealed a constant ("prevailing") component upon which the tides are superimposed, see Figure 4. The prevailing component achieves world-wide transport from the summer to the winter hemisphere (Figure 7) so that it exhibits a strong seasonal influence (Figure 8).

All together, the height vs. latitude patterns are quite involved (Figure 9). An unexpected detection was a quasi-periodic variation with about two days period (Figure 10); it is explained as a higher order zonal wave (Kalchenko, 1987).

Tidal winds are produced by solar heating on the one hand, and by upward energy transfer from perturbations at tropo- and stratospheric levels on the other (see, e.g., Forbes, 1987).

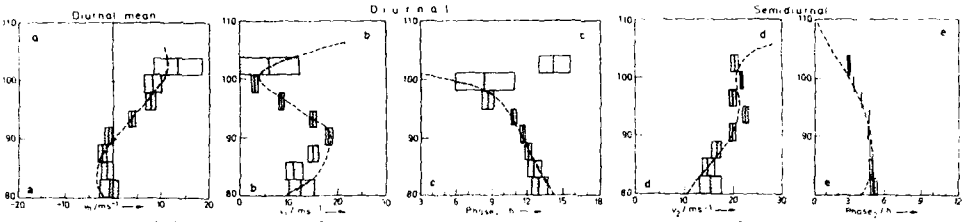



Figure 4. Height profiles of tidal components: prevailing - diurnal - semi-diurnal /velocity at left, phase at right hand side/. (a) Average for notherly wind (southward wind vector). (b), (c) Amplitude of rotating wind (phase = time of northward maximum). Reduced from a set of 24.5oo meteor echoes at Kyoto (Japan) between 2o July and 7 August 1978 (Aso et al., 198o).

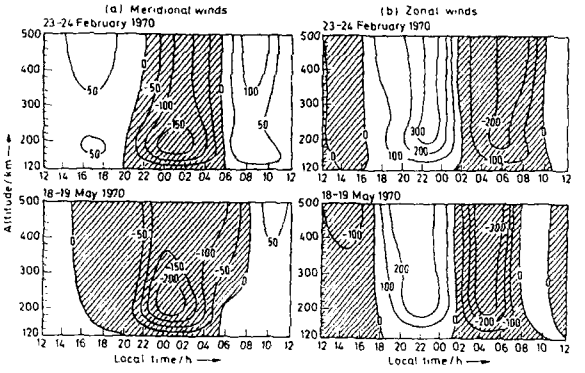



Figure 5. Typical seasonal height profiles of neutral winds inferred from semi-empirical theory fitted to (incoherent scatter) plasma drift measurements at 43°N (Roble et al., 1977).

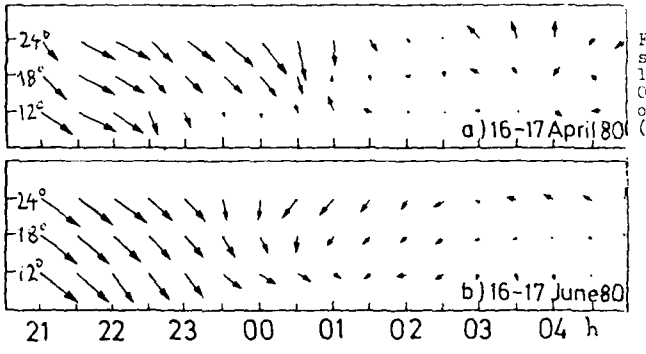



Figure 6. Horizontal thermospheric winds in a latitude vs. longitude (local time) frame. Obtained (for night hours) by optical (airglow) observations (Burnside et al., 1981).

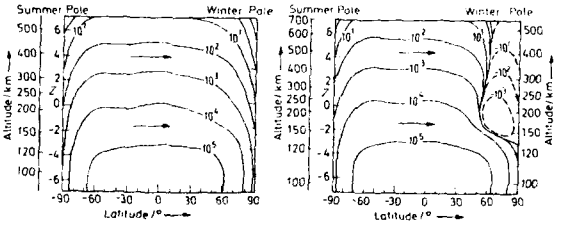



Figure 7. Maps of meridional average transport (zonal means) showing total air transport (in kg/s) down to the indicated altitude. Left hand map is for low, that at right hand for high solar activity; the reverse cell is due to auroral zone heating. Semi-empirical computation (as in Figure 5) by Roble et al., (1977).

It is worth noting that the U.S. National Center for Atmospheric Research runs a computer program producing a semi-theoretical model which needs being fed with measured or assumed data from a small number of stations; see Figure 11 (Salah et al., 1987).

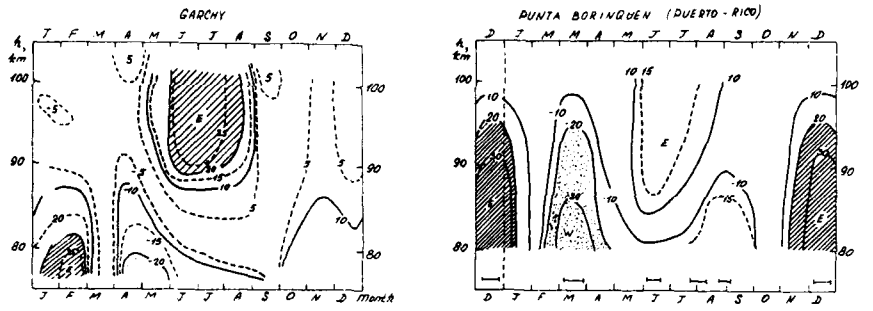



Figure 8. Prevailing wind as function of height and season after 1977 to 1979 (meteor radar) observations in France and Puerto Rico (Massebeuf et al., 1981).

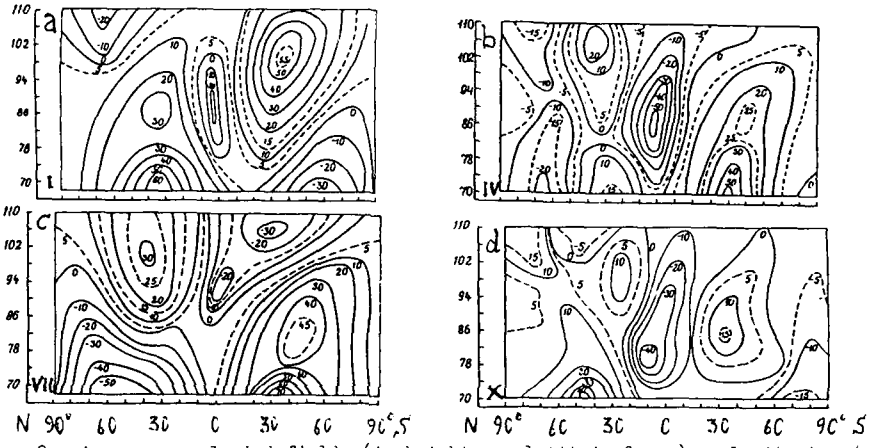



Figure 9. Average zonal wind fields (in height vs. latitude frame): velocity in m/s, westerly positive. Results from meteor Radar and partial reflection observations at 27 stations in both hemispheres. (a) January, (b) April, (c) July, (d) October. (Fortnyagin, 1987).

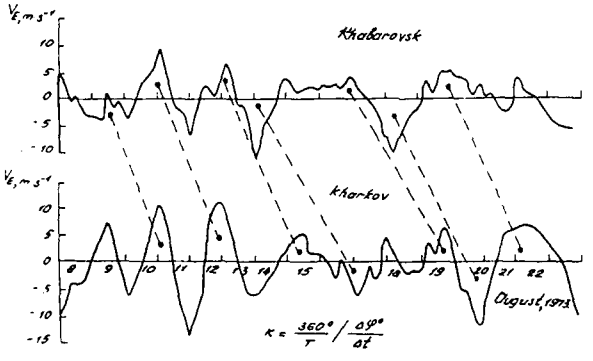



Figure 1o. Quasi-two-days variations of the zonal velocity at around 1oo km of altitude. Simultaneous observations at Kharkov and Khabarovsk, different in longitude by 98°. /The phase lag of 62 to 65 h might be due to a zonal wave number 5/. (Kalchenko, 1987).

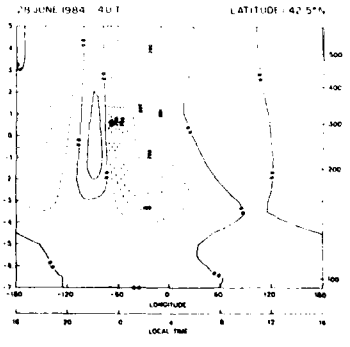



Figure 11. Meridional wind component (northward positive) in a height vs. longitude frame. (Fixed latitude and UT). After insertion of measured data computed with the semi-empirical NCAR "Thermospheric General Circulation Model". (Salah et al., 1987).

## 2.3 Gravity waves

In tides, irrespective of the relevant driving cause (solar heating, gravity forces), buyoncy is the "counter-force", pushing the system back towards the equilibrium constellation. (This differs from acoustic waves, in which wave pressure plays that role). So, tides are gravity waves on a global scale.

In fact, there appear two separate mode ranges in the well-known wave diagram of Figure 12. Tidal waves correspond to extremely low frequencies; they are "located" at the bottom line of the diagram. The range of gravity waves extends, however, far beyond these frequencies, up to the Brunt-Vaisälä frequency, i.e. down to periods of about 5 min. Gravity waves of so short period can, of course, not extend all around the globe but must be localized. Excited by different mechanisms, they are a permanent feature of the upper atmosphere, not only as a disturbance of the steady state but also by affecting the circulation in the middle atmosphere and lower thermosphere. Energy transport by such waves, upwards in most cases, is not negligible (Klostermeyer, 1973).

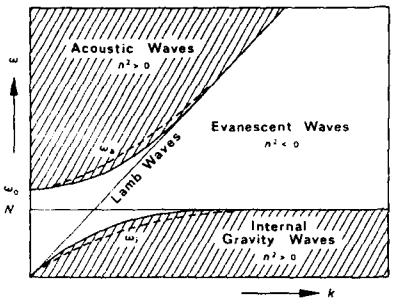



Figure 12. Frequency (ordinate) vs. wave number (abscissa) dispersion diagram showing the ranges of acoustic and of gravity waves. After Tolstoy (1963). (See also Gille, 1968).

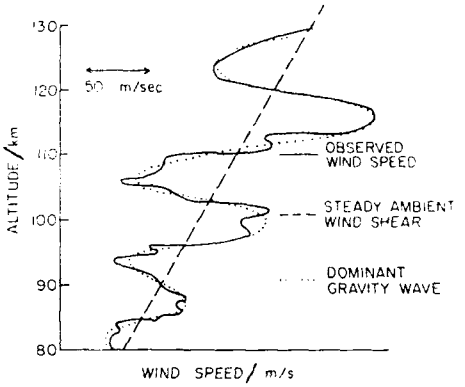



Figure 13. Gravity wave: schematic wind velocity profile (Midgley and Liewohn, 1966).

Table 1 shows typical parameter values: period, wind velocity amplitude, vertical and horizontal wavelength (see Figure 13).

TABLE 1: TYPICAL CHARACTERISTICS OF ATMOSPHERIC GRAVITY WAVES

| Identification | Period/h | Velocity/$ms^{-1}$ | $\lambda$ vert/km | $\lambda$ hor/km |
|---|---|---|---|---|
| (a) all | 0.5 .. 6 | 6 .. 7o | 1o .. 3o | 1oo .. 8oo |
| (b) large scale | >1 | >5oo | | |
| (c) medium scale | 0.2 .. 1 | 9o ..4oo | | |

(a) meteor Radar (Kalchenko et al., 1985; Gavrilov, 1987);
(b) different techniques (Hunsucker et al., 1986).

Different techniques show different results because the range of detectable periods depends on the technique. Short periods are difficult to identify, in particular in meteor Radar and in total electron content measurements. The shortest periods can probably be seen with continuous ionospheric sounding (Nakata et al., 1953), see Figure 14. Quite short periods, down to about 1o min, are seen by day; the oscillations are slower and stronger by night. This is clearly appearing from Figure 15.

Such disturbances were first seen in quick sequences of ionograms (Figure 16) as deformations of the echo trace starting at the top (Eibl et al., 1959). Though, apparently, coming down, they do not originate at the top of the ionosphere as first was thought.

Hines (1960) established that these "travelling disturbances" are due to almost horizontally propagating "internal" atmospheric gravity waves. The wave fronts are forward inclined what explains the observed facts. Since the vertical wavelength is always much smaller than the horizontal one, the phase propagation vector is almost horizontal with a small, mostly downward, vertical component. Energy flux, however, is usually upwards.

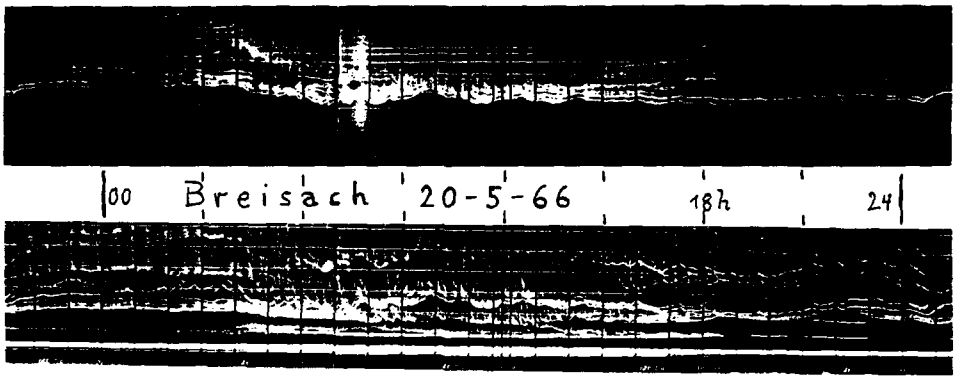



Figure 14. Continuous ionosonde records during a period of almost 24 h. MUF on top, virtual heights below. Note the short period fluctuations during daylight hours. (Courtesy: K. Bibl).

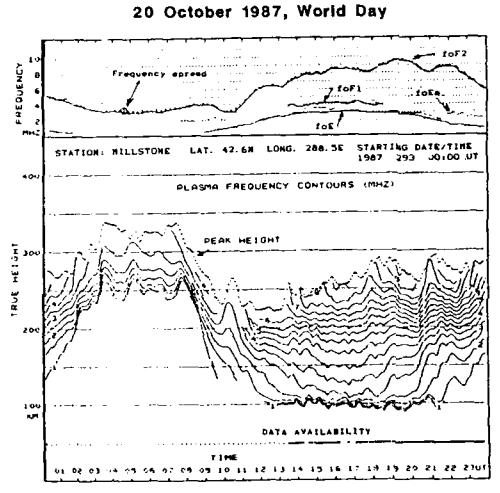



Figure 15. True height contours of plasma frequency over one day. Digital ionosonde (with automatic true height data edition) of Lowell University, Ma. (Courtesy: K. Bibl and R.W. Reinisch).

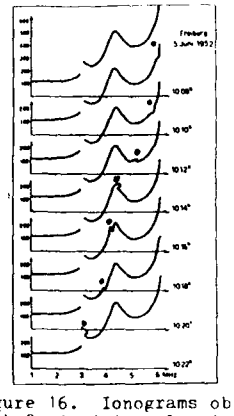


Figure 16. Ionograms obtained with 2 min intervals. A deformation starting at 10h10 moves down the trace (Bibl et al., 1959).

In Figure 13 we have shown a very clear gravity wave in the lower thermosphere. Such "ideal" conditions are rare. Most often appears a mixture of several waves. Figure 17 gives an example in which frequency analysis was applied to meteor radar data (Gavrilov, 1987).

Though beacon satellite observations of the total electron content give an integral value over a long path, even in these observations gravity waves could be detected. Figure 18 shows a clear case in which an unique event was the cause of the observed wave. A large statistics of wave periods observed at a high-latitude site /using simultaneous emissions of two satellites/ is shown by Figure 19. The range extends from about 0.2 to 1.3 h, but 75% of all data are inside 25 to 50 min.

Gravity wave effects are also apparent in the characteristics of the $E_s$-layer ("sporadic E") and in airglow brightness data (Charadze et al., 1987).

As for the sources of atmospheric gravity waves, it is certain that there are many. Hunsucker (1985; ... et al. 1986) feels that "large-scale" waves (Table 1) are produced in the polar regions and propagate from there, mainly toward lower latitudes as was earlier stated by different authors. The wave (wind) amplitudes indicated in Table 1 are so large that non-linear phenomena must be expected, and instabilities can be created (Mied, 1976; Klostermeyer, 1987).

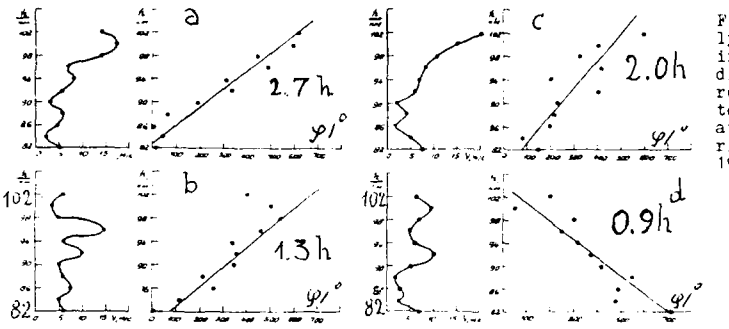



Figure 17. Height analysis of one observed internal gravity wave distinguishing different periods (from 0.9 to 2.7 h). Amplitudes at left, phases at right hand. (Gavrilov, 1987).

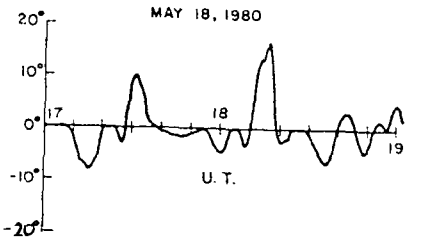



Figure 18. Faraday angle variations observed at Urbana (Ill.) after the St.-Helen eruption (18 May 1980, [illegible] UT). Observed quasi-period: [illegible] h (Liu et al., 1982).

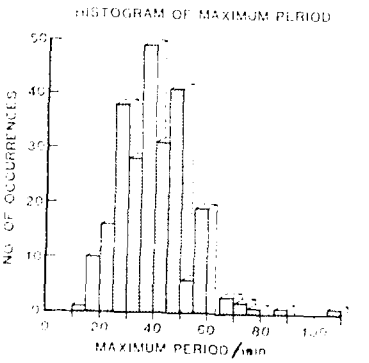



Figure 19. Wave period statistics from electron content data obtained in slants with two geostationary satellites. (Bunchuker et al., 1986).

Different from that, medium-scale waves are found to have no preference for a certain direction. Most of these are thought to come from tropo-stratospheric sources, in particular the jet-stream (Bertin et al., 1975). Another presumable source is the dynamo-region, near the E-peak (Klostermeyer, 1969). It is worth noting, that - as a consequence of atmospheric sedimentation - the amplitude of a gravity wave should considerably increase when propagating upwards. Therefore, rather weak sources at a lower level can produce large effects higher-up, except when the wave is (partially) reflected inbetween (Hines, 1968).

Gravity waves can produce rather confusing instantaneous wind patterns, particularly in the E-region, see Figure 20. When looking for systematic effects in ionospheric dynamics, these waves are classified as "noise".

## 3. REACTION TO EXTERNAL ELECTRIC FIELD

In "cooperation" with the terrestrial magnetic field, an external electric field gives rise to a cross-field drift of velocity, $\underline{E} \times \underline{B}/B^2$, when collisions can be neglected (as in the upper thermosphere), the plasma as whole moves in cross-field direction.

As a consequence of the anisotropy of the plasma conductivity the field-aligned and cross-field components of an applied external electric field behave differently. The field-aligned component - due to good conductivity in that direction - does almost break down, (in-situ measurements ended-up with quite small fields). On the other hand, since the cross-wise conductivity is very small, the cross-component is preserved and is transferred along the well conducting "magnetic" so that these become potential lines of the field. Only at levels where collisions are more frequent, the medium does reach some cross conductivity by which the cross-field is diminished. This happens near the base of the ionosphere.

### 3.1 Fountain Effect

[illegible] (up and down) motion of the conducting ionospheric layers induce strong zonal currents. These are the cause of the regular diurnal variations of the magnetic field. As a consequence of upwards increasing "insulation" between fieldlines, the currents are strongest in the lower thermosphere where frequent collisions make the cross-field conductivity large enough (e.g., Volland, 1981).

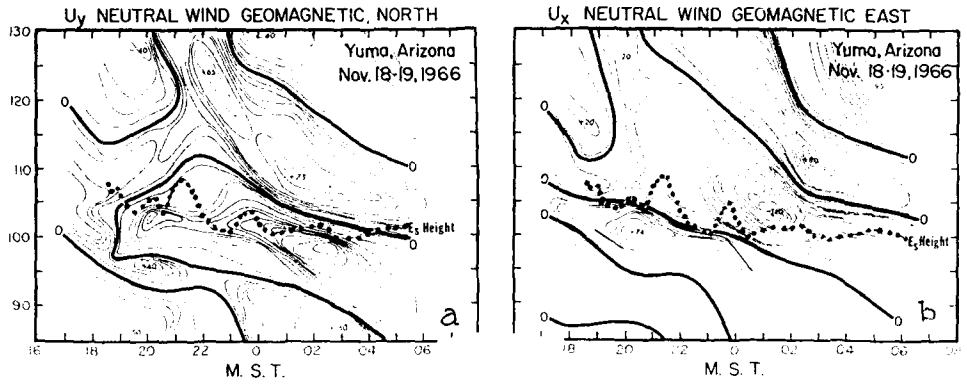



Figure 20. Neutral wind instantaneous contours derived from optical observation of seven successive (gun-launched) metal vapour releases. (a) Meridional component (geomagn. North); (b) Zonal component (geomagn. East). Bold dots identify height variation of a single sporadic $E_s$-layer observed by ionosonde. /The probable connection with wind shears is not further discussed here/. (Wright, 1966).

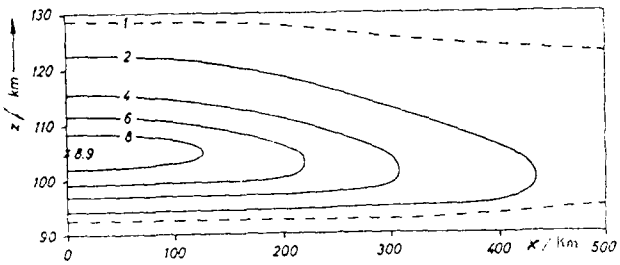



Figure 21. Zonal equatorial electrojet, meridional section, true magnetic equator at left hand (x = 0). Isolines of current density (in $10^{-6}$ A $m^{-2}$). (Wright, 1966).

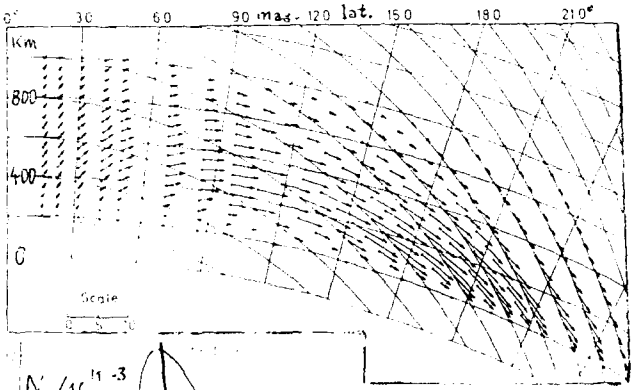



Figure 22. Plasma drift velocity in a meridional section, magnetic equator at left hand. Computed by combined action of electrodynamic lift and diffusion. /scale unit: $10^{15}$ $m^{-2}$ $s^{-1}$/. (Hanson et al., 1966).

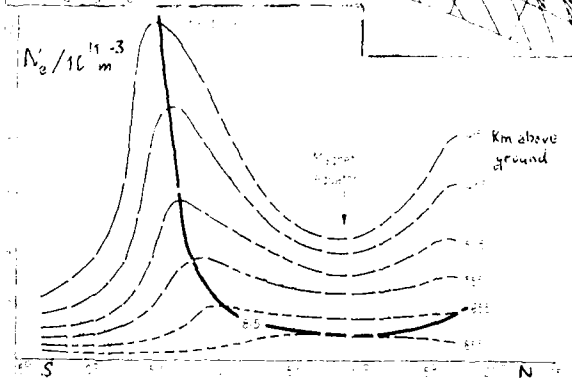



Figure 23. Contours of electron density in the equatorial anomaly after topside sounder-observations near Singapore. (King et al., 1963).

As shown by Figure 21, the largest current density is reached at the magnetic equator; there is the center of the "equatorial electrojet". Since the magnetic field is horizontal, the plasma drift is vertically upwards during day light hours. This "fountain phenomenon" is regularly occurring in the equtorial belt (see Forbes, 1981). The lifted plasma flows down along the fieldlines and so increases the plasma density north and south of the equator, while it is decreased (and the peak raised upwards) at the dip equator itself. The dynamo field disappears by night; the plasma then comes down and moves towards the equator, supported by a neutral equatorward wind.

Figure 22 presents computed results of plasma motion in the "fountain" (at left) and away from it, while Figure 23 shows contours of plasma density observed by topside sounding (King et al., 1963). Almost all individual days exhibit the two maxima in the meridional noon plane. However, their position in latitude is somewhat variable and the meridional gradients are considerably steeper for individual days than is appearing on monthly median maps (Vila, 1971).

## 3.2 Fields Transferred from the Topside

The conditions are quite different in the polar caps because, there, the magnetic field lines are open such that transfer of plasma and electric fields occurs from the magnetospheric tail and is most important in determining the dynamics of the polar ionosphere. Figure 24 shows F-region drift measurements made at a site near the center of that cap: plasma flux is continuously anti-sunward such that the direction rotates by 360° in 24 h. In magnetic coordinates the rotation is approximately linear in time as the site rotates around the magnetic pole (Reinisch, Buchau et al., 1967). Such cross-field motion is provoked by an electric field inferred from the magnetospheric tail.

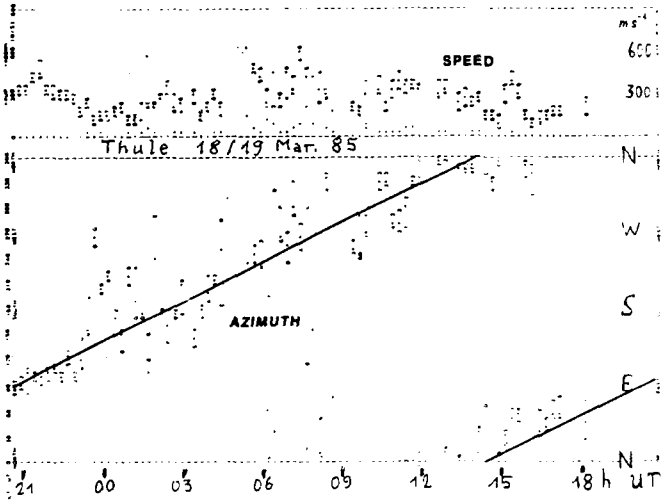



Figure 24. F-region drift measurements 18/19 March 1965 at Thule (77.5°N, 291°E), speed on top, direction below. /Spaced antennas Doppler technique/. (Reinisch, Buchau et al., 1967).

In the auroral zones, conditions are more involved and less regular. Details shall not be discussed here, but again electric fields are of major importance.

The different mechanisms and the relevant transport equations have been compiled by Schunk (1975).

## 4. PLASMA DRIFT MEASUREMENTS

While the individual ions of a plasma have quite important thermal velocities, their overall average (or "bulk") velocity is considerably smaller. It is the latter which is intended when trying to measure "plasma drifts". A direct measurement of the statistical average of individual ions was, for example, achieved aboard a spinning spacecraft, the Venus-Orbiter (Knudsen et al., 1979). This, however, is an exception. Most methods depend on "tracers", the drift of which can more accurately and easily be determined. The question arises, whether the motion of a tracer is identical with the plasma bulk velocity.

A most efficient tracer method applied to neutral air motion is the Radar pursuit of the ionized traces of meteors. This method has become the most widely used for determining neutral winds in the lower thermosphere (see Figures 4, 8 to 10, 17). Unfortunately, there are no comparable tracers of independent origin available for measuring the motion of ionospheric plasma as such; internal tracers only can be used to this end.

## 4.1 Doppler Methods

Radio waves echoes obtained from the ionosphere by any technique must be Doppler-shifted in case the origin of the return is moving in such way that its effective distance varies. The speed of the motion along the "effective sounding direction" can be deduced from that frequency shift.

### Echo Doppler Sounding

When the (rf or if) echo signal is phase-compared with the transmitted one, the frequency difference appears as a low frequency beat. Ranzi (1933) was first in applying this technique which was perfected later by Bindlay (1951). With modern digital techniques the beat frequency can be obtained with high precision in digital form, see Section 4.3 (Heinisch, Bibl et al., 1987). Notwithstanding whether one uses a monostatic system (vertical sounding) or a bistatic one (oblique sounding) the beat depends on the vertical displacement of the reflecting surface, i.e. that on which the critical electron density is first reached. Sunrise and sunset effects are clearly seen but also systematic height variations, and gravity waves as well. Note that the decisive parameter is phase (not group) velocity and that it depends seriously on the local electron density, i.e. it is height dependent. Therefore, a quantitative interpretation of the observed data is not easy.

### Incoherent Scatter Technique

A very weak backscatter signal is obtained when radio-sounding on a frequency high above the ionospheric plasma frequencies. By analyzing the spectrum of the return signal one obtains the ion temperature from the spectral width, while the average motion is determined from the shift of the return center frequency against the transmitted spectral line. The method does not measure the motion of individual ions; the return signal is due to small-scale fluctuations of the local plasma density as are brought about by statistical fluctuations in the thermal motion. These, and their behavior in space and time, are correctly described by Fourier analysis, namely as a stochastic set of waves in the ion gas. The short wavelengths corresponding to thermal motions fall into the range of sound waves (see Figure 12). Their range is larger than the sensibility range of the radar, so that this latter selects a smaller range inside which the (projection of the) sound wavelength is not too much different from the vhf wavelength of the radar. Therefore, one observes in fact a section of the thermal spectrum. Since there is no reason against the assumption that the whole set of soundwaves should be isometric, it is justified to assume that their (vectorial) average equals the bulk velocity of the ions.

Mono- and bistatic incoherent scatter installations can only observe in one direction, while multistatic installations can determine a spatial drift vector. Drift data can now be obtained in quite short time intervals so that short term variations can be seen; see Figure 25 in which, apart from quick changes, the M2 tide and its height-dependent phase do quite clearly appear.

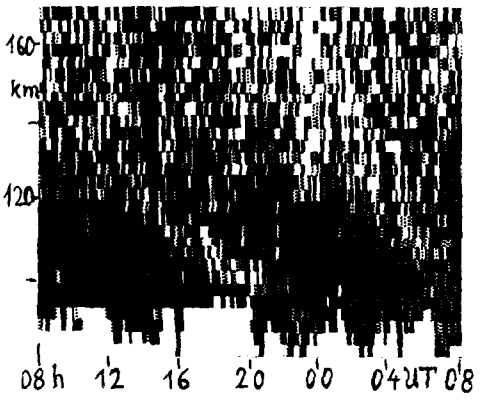



Figure 25. Field-aligned ion bulk velocity observed by incoherent scatter technique during a 24 h interval. Darkest fields correspond to about 150 m/s northward drift, blanck fields to the same velocity southward. (EISCAT Annual Report 1986).

Whenever observing along the field line (as done in Figure 25), the combined effects of wind-induced motion and parallel electric field are seen; only when this latter is negligible, the neutral (horizontal) wind can be correctly determined from incoherent scatter observations along the magnetic field (Miller et al., 1987).

On the other hand, cross-field motions of the plasma can be obtained with the same device by beam-swinging. Figure 26 shows quick variations (as often found in the auroral zone) in the 170 to 550 km height range. Cross-field motions must be caused by an electric field ($\underline{E} \times \underline{B}/B^2$ drift) such that such fields can correctly be measured by devices of that kind.

Incoherent scatter sounding at oblique incidence against the magnetic field needs careful analysis of the different causes (Calah, Luboin et al., 1987).

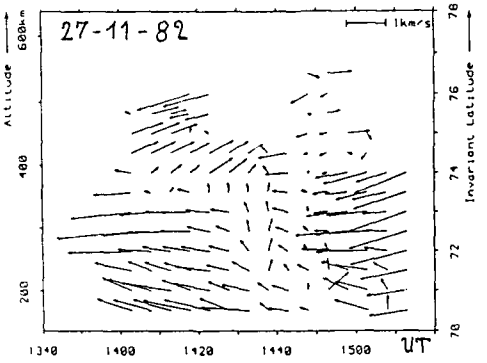



Figure 26. Plasma flow vectors normal to the geomagnetic field direction obtained by beamswinging. (Jalonen, 1986 after van Eyken et al., 1984).

4.2 Direction Finding Methods

These devices apply triangulation methods, usually with multistatic coherent radar installations. These are applied on either vhf frequencies ("coherent back scatter") to the whole ionospheric height range, or, in the subpeak range, hf frquencies using total or partial reflections. With total reflections, deformations of the (plasma) equidensity surfaces are used as tracers while, for the other techniques, natural phenomena provoking steep plasma gradients are at the origin of the observed returns. The situation is similar to that with wind determination from drifting clouds in satellite meteorology. As for these, local changes occurring to the tracer phenomenon can seriously influence the result. Only if this danger is eliminated, the latter can be taken for granted.

For a long period, the following four techniques were distinguished (Rawer, 1968):

L1 : Intercomparison of fading signals at three or more antennae spaced a few wavelengths apart;

L2 : Radio observations of drifting meteor trails;

L3 : Radio star scintillation with three or more antennae spaced many wavelengths apart;

L4 : Observations of characteristic reflection features at widely spaced sites.

The L2 method depends on external tracers and is, therefore, well indicated for measuring the neutral wind. Since in the (lower thermospheric) height range, where meteors are observed, neutral and plasma motions are coupled by collisions, L2-data have often been used for checking other methods. Good agreement with L1-data was found in the above mentionned height range (Müller, 1968).

Method L3 differs from the others as it observes penetrating, not reflected waves. Applied on vhf-frequencies, the observed fadings are mainly due to refracting irregularities near and above the peak of the ionosphere. The method was only rarely used. In recent time, coherent radio beacons aboard geostationary satellites are used as sources - Figures 18 and 19 above were obtained with this technique.

Method L4 has also only occasionally been used by comparing simultaneous records from ionospheric stations distant by 100..300 km. Trace deformations, that are coordinated and can be identified over such distance, are necessarily large scale irregularities or tilts.

As a general rule, the extension of "usable" irregularities depends on the basis length used in the experiment: larger structures are seen with larger bases (Rawer, 1963). So, according to its extension, a device may detect structures ranging from about 10 m to more than 100 km, and periods of the order of less than 1 to a few min. As the (Brunt-Vaisälä) cut-off of ionospheric gravity waves is found at periods between about 3 and 6 min, some of the irregularities could propagate as gravity waves, other must fall into the evanscent range of Figure 12, or even into the sound wave range. The latter two classes should be transported by the (local) wind. However, longer period gravity waves are regularly present and the relevant motions are seen as "instantaneous local wind" - which is measured. If the directions were equally distributed, then a large scale bulk velocity could be detected as a statistical mean - otherwise the average designates the preferred direction of the gravity waves. This should depend on their origin.

Dispersion must exist such that different phase velocities are to be expected. Also, the life-time of irregularities should depend on their size, smaller ones having shorter life. Therefore, large-base arrangements only can see the fastest moving irregularities: with longer bases one must find higher velocities.

Another important problem is connected with the shape of the irregularities. The early reduction schemes assumed isometric patterns. We have, however, learned that the

irregularities should be elongated in the magnetic field direction, (e.g. Bellchambers et al., 1965), except for the lowest thermosphere, where they should be horizontally stratified (as a result of sedimentation in the gravity field). So, we have to accept that the patterns are anisometric. Figure 27 (below) shows examples.

Comparison of different techniques and reduction methods has shown that the drift direction is much better measurable than are absolute speeds. (See, e.g. /9/, also Jacobs, 1968; Müller, 1968; Wright, 1968).

Different Reduction Methods

When applying triangulation methods, one compares records of hf fieldstrength (or any other wave characteristic) that are taken at different sites. Quite different reduction methods have been proposed and are used for deducing a motion vector.

In one method, individual features are identified on the records and time differences between their appearance at the different sites are determined. With three stations, assuming horizontal motion, one can deduce a drift vector. This is the method called "similar fades". It must be applied to many fades: usually about 2o during an observation period of a few min. The statistical average of these is taken as plasma bulk motion vector. The individual determinations show considerable dispersion. In fact, when assuming a wave-like propagation, the relevant fronts should be moving across the antennae array. The measured velocity is, necessarily, perpendicular to these fronts "of constant phase". Their traces on ground should form a system of equidistant, parallel lines. Observations with a 6-antennae array have, however, shown that this is rather often not true.

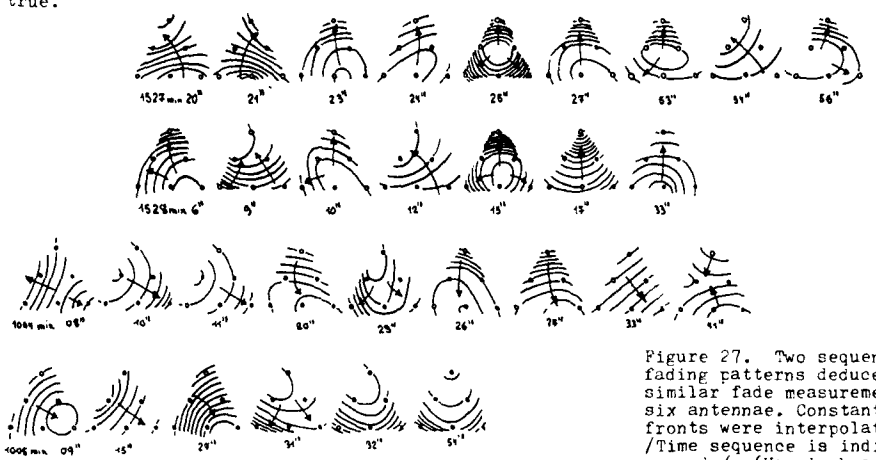


Figure 27. Two sequences of fading patterns deduced from similar fade measurements with six antennae. Constant phase fronts were interpolated. /Time sequence is indicated in seconds/. (Harnischmacher, 1968).

It appears from Figure 27 that "exploding" or "imploding" patterns do sometimes occur, in which case the individually determined vector cannot be significant at all. One should not forget that the interference pattern seen at ground often is due to several contributing returns.

Other methods use the complete records, determining cross-correlation functions between pairs of these. Figure 28 may help to understand the basic idea (assuming isometry and that no vertical motions must be considered). The displacement in time of the points of maximum cross-correlation (in a temporal diagram) can be used to define a (horizontal) characteristic velocity vector. If the observed patterns are isometric, this is the intended "drift velocity" (using "maximum correlation" as a kind of "tracer"). Unfortunately, isometry is an unrealistic assumption: the spatial correlation patterns have elliptical shape and this feature must be taken into account (Phillips et al., 1955; Briggs, 1957).

When the above made assumptions were valid, data from three positions, by the relevant auto- and cross-correlation functions, could be used for determining all unknowns. (A very helpful formulation was given by Keneshea et al., 1965; see also Pfister and Libl, 1972).

Unfortunately, the patterns are not isometric, nor is it true that vertical motions are negligible (see Figure 38 below). A three-dimensional analysis was first made by Gusev et al. (1960) and later by Duchy (1966), who attacked the problem in a quite general way. He concluded that information from any array situated in one plane only is insufficient. In order to resolve after all unknowns, one must at least dispose of one more measuring position situated above the ground plane. With other words: a three-dimensional problem cannot be fully investigated with a two-dimensional measuring array.

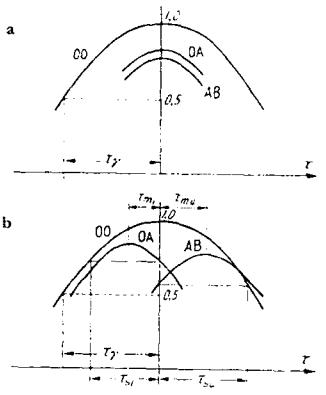



Figure 28. Correlation analysis to a drift measuring array consisting of one center position, O, and two side positions, A, B. The correlation functions show how the correlation coefficient varies with the time shift, $\tau$, by which the two records are displaced against each other. OO is autocorrelation for the center, OA and OB are cross correlations with the side positions.
(a) Only unordered (isometric) motions, bulk velocity = 0 /or: the measuring device moves with the bulk velocity/.
(b) Realistic situation with not vanishing bulk velocity. (Rawer and Suchy, 1967).

Fading Drift Measurements (Method D1)

Natural deformations of the (otherwise plane) surfaces of constant plasma density are provoked by different phenomena, in first line by gravity waves. When applying classical pulse sounding in the hf range, echoes from two or even more reflecting ranges in the ionosphere do arrive at the receiving antenna, so that a fading pattern is produced by superposition. These fieldstrength variations are recorded in an array of, usually three, spaced antennae (Krautkrämer, 1943/1950).

An (apparent) velocity vector is either obtained by evaluating a certain number of typical features (method of "similar fades"), or by a correlation method (see the last paragraph). Both methods were compared at different places with the general result that, in the statistics, the directions do agree within about $\pm 30^0$ (Bellchambers et al., 1965). As for the speeds, the similar fade evaluation gives greater values (by almost a factor of 2) than the correlation method pretends. However, optical observations of releases gave even higher values than the similar fade method achieves. Optical data probably indicate the true neutral air wind. The discrepancy between both evaluation methods was explained by Sprenger et al (1969). These authors could show that large-scale (very low frequency) deformations, of the order of 100 km (horizontal) wavelength, seriously influence the results of correlation analyses. When filtering away these frequencies from the records, comparable speed values were obtained with both methods. Therefore, most of the published data that were obtained by correlation analysis, should be corrected in speed.

Because they are made at a fixed radio frequency (potentially switched between day and night), D1-measurements are restricted to a few limited height ranges, viz.
85 .. 100 km for lf sounding,
95 .. 120 km for hf sounding, gate on "E",
170 .. 250 km for hf sounding, gate on "F".
It is a certain disadvantage that the reflecting level is not constant but has a diurnal variation inside these height ranges. Thus, caution is needed when comparing data obtained by day and by night.

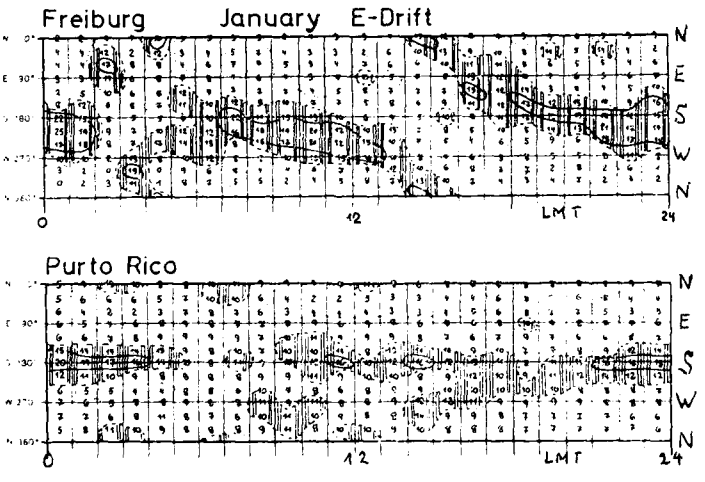



Figure 29. D1-drift E-region: (hourly) percentage statistical distribution (small digits) January average over about 6 years. Percentages above 9 are hatched. (Harnischmacher, 1966).

D1-measurements were quite frequently made in the Eastern hemisphere. The data obtained constitute the numerically most important thesaurus of drift measurements. It is important to note, that the systematic effects, which have been found, are almost drowned in "noise"; in fact, the instantaneously measured vectors reproduce local, small-scale motions largely due to gravity waves. This can be seen from Figures 29 and 30. The latter shows that it is important to make enough measurements, hour by hour if possible, in order to identify regular features. These can only be extracted by statistical methods.

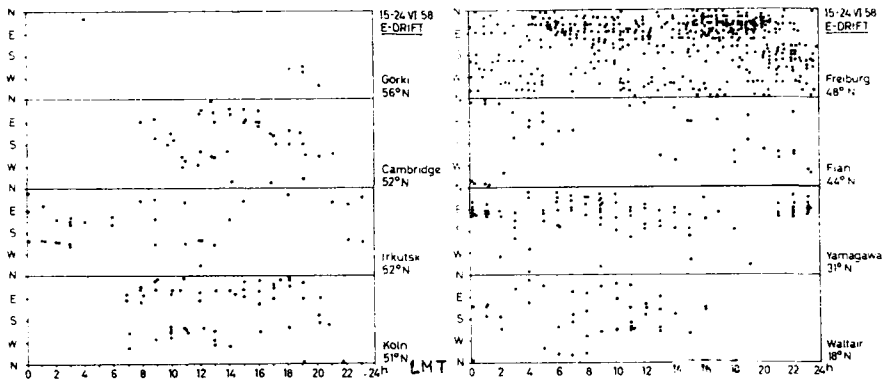



Figure 30. Intercomaprison of E-drift directions measured by method D1 at 8 stations in the period 15 to 24 June 1958 (Rawer, 1965).

Probability difference diagrams are obtained by deducting from the observed percentage probability values (e.g. Figure 29) the "equal probability distribution" - (i.e. 8.3% in our arrangement; this brings the hourly sum to zero): Figure 31.

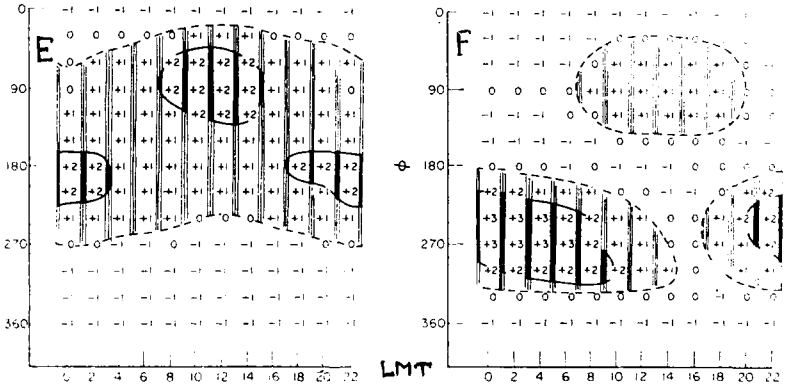



Figure 31. Probability difference diagrams showing the dominant component's average diurnal variation. All summer observations of many years made at Freiburg. (Harnischmacher et al., 1968).

Pertinent features can be "extracted" by making "sliding averages" over several fields. According to the chosen direction (in the diagram), a certain feature is amplified: horizontal averaging to get the "dominant components" (which are different for day and night: Figure 32); oblique averaging - according to the chosen averaging scheme - to get the 24 h or 12 h (rotating) solar tidal components. /Figures 31 through 35 refer to all drift directions measured at Freiburg during almost one decade. E-echo results are always at the left hand, F-region ones at the right hand side/.

Figure 33 is the seasonal direction statistics. Figures 34 and 35 show the 24 h and 12 h solar tides. The lunar effect was extracted in Figure 36.

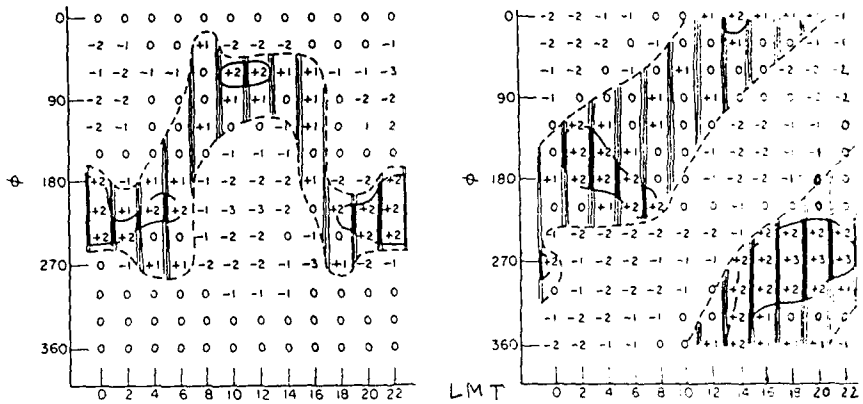



Figure 32. Probability difference diagrams after amplification of the dominant component (by slidingly averaging). (Harnischmacher et al., 1968).

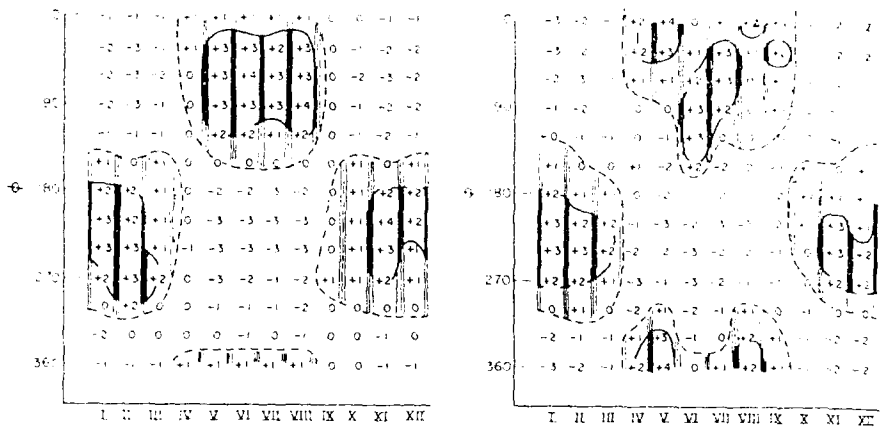



Figure 33. Probability difference diagrams showing the average seasonal direction statistics (obtained by orizontally averaging over 24 h in each monthly diagram). (Harnischmacher et al., 1968).

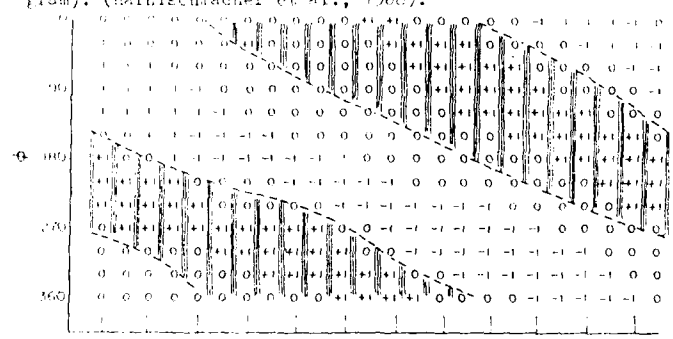



Figure 34. Probability difference diagrams after obliquely averaging so as to extract the solar 24 h component. (Harnischmacher et al., 1968).

A comparison of lf measurement (L and lower E-region) with meteor radar data is shown in Figure 37. Mainly zonal drifts are observed in this height range with day/night and equinoctical reversals (Greenger, 1960).

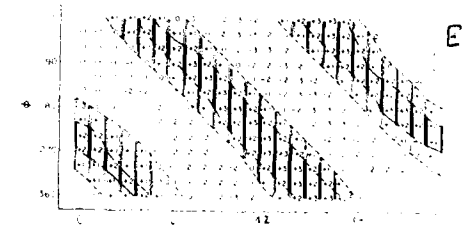



Figure 35. Probability difference diagrams after obliquely averaging so as to extract the solar 12 h component. (Harnischmacher et al., 1968).

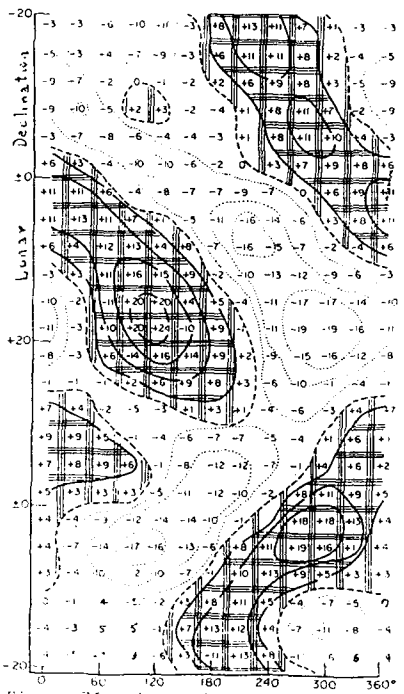



Figure 36. Probability difference diagram showing the lunar declination (ordinate) influence on the drift direction (abscissa) obtained by grouping all observations made at Freiburg after the lunar declination. E-echoes. (Harnischmacher et al., 1968).

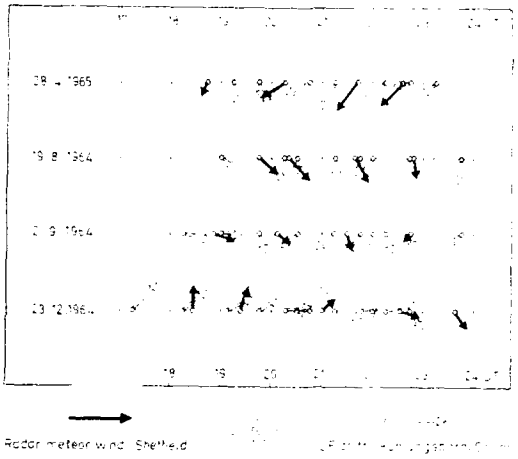


Figure 37. Synoptic comparison of Lf drift measurements made (by method LI) at Kühlungsborn and at Collm (GLR), compared with meteor radar measurements made at Sheffield (U.S.A.). (Greiner and Schminder, 1968).

## 4.3 Combined Doppler and Direction Finding

The LI-method utilizes the fact that the equidensity surfaces in the ionosphere are not perfectly plane. This is known since long time; Bramley et al. (1951) found - even under quiet conditions - the mirror point to move irregularly inside a circle of about 10 km radius. With improved accuracy of the measuring devices, Essex et al. (1968) investigated the correlation between adjacent sounding frequencies and found a correlation width of only 20 .. 40 kHz; this says, that the decisive irregularities must be quite near to the reflecting level.

Nowadays, with modern digital electronics, such investigations can be made in real time and with very high accuracy. As mentionned under 4.1 above, digital ionosondes (Pitl et al., 1978, Reinisch, 1986) regularly can measure the Doppler shift (order: Hz) of the received echoes. This technique, recently, was combined (Reinisch et al., 1986) with the spaced antennae direction finding technique (in a fixed frequency "drift mode"). Time series are received at each antenna of an array of four. Real-time Fourier-transformation results in four complex spectra (resolution 0.125 Hz). The angle of arrival for each spectral component is determined by cross-correlating these complexe spectra.

With such techniwue it could be shown that, usually, a large number of return sources (more than 1o) coexist - not as a stable configuration but with quick changes. One needs - similarly to known experience with D1 - about 15 min for obtaining significant averages so as to reach a consistent picture of the systematic, larger scale variations (see Figure 38). Another interesting result of that technique, obtained at a site in the polar cap, was shown on Figure 24 above.

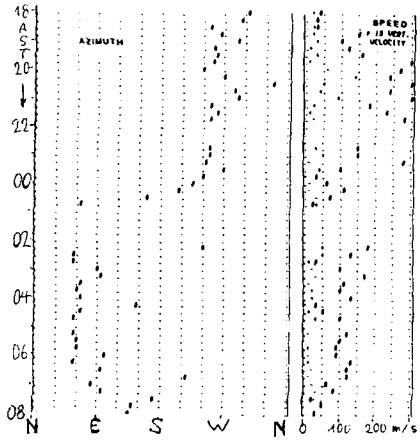



Figure 38. Variation of F-region drift measured with a digisonde and a 4-antennae-array during the night 14/15 Jan. 1983 at Goose Bay, Labrador. Azimuth at left, speed at right hand. (Reinisch et al., 1987).

In the future, this very promising new technique - when applied at more stations - may shed much more light on ionospheric motions so that, finally, we may get a new, empirically based picture of ionospheric dynamics.

## 4.4 Drift Data in the IRI

The "International Reference Ionosphere" (IRI) /an URSI-COSPAR project/ aims at a representative, empirical world-wide description of (vertical) profiles of the most important plasma parameters: density, temperatures, chemical composition. Since a few years /on the initiative of K.S. Kazimirovsky/ a tentative description of ionospheric drifts is additionally given. Basic data are mainly from D1- and D3-measurements. They describe the *apparent motion of medium scale* (1o .. 1ooo m) irregularities of plasma density. The 1957 to 1979 data of stations situated between 7° and 71°N (7.5° to 64° geomagnetic) and 0° and 131°E in longitude. Some $10^5$ individual measurements were separated after regions E (9o to 135 km virtual height) and F ( > 135 km), and in seasonal and solar activity classes. For six key latitudes a Fourier-development up to order 3 was made, separately for the zonal and meridional component as function of local time.

A preliminary world-wide description was made with the so found coefficients, and spline interpolation. The so obtained maps have latitude and local time as coordinates. Figure 39 shows an example.

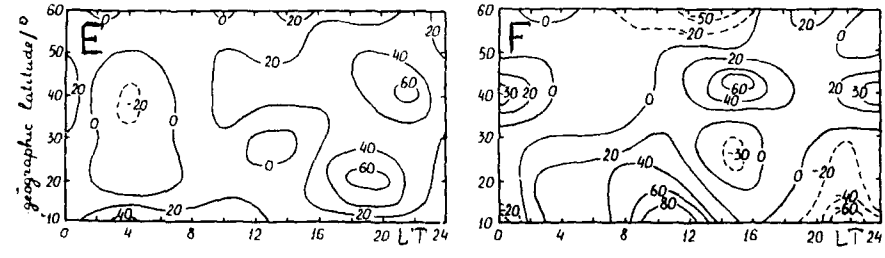



Figure 39. Preliminary drift model: global distributions of the zonal velocity component (unit m/s, E positive) in summer. Left: E-region, right: F-region. (Kazimirovsky et al., 1988).

Though there is little doubt that this can only be seen as a first attempt, we feel that it could become very important to aim at a fully empirical description of ionospheric drifts - a task for future work in the field.

REFERENCES


/1/ COSPAR (Committee on SPACE Research of ICSU), "CIRA 1972, COSPAR International Reference Atmosphere 1972", Berlin, Akademie-Verlag, 1972.

/2/ RAWER, K. (ed.), "Winds and Turbulence in Stratosphere, Mesosphere and Ionosphere", Amsterdam, North-Holland Publ. Comp., and New York, Academic Press, 1968.

/3/ ROPER, R.G. "Handbook for MAP, Volume 25", Urbana, SCOSTEP Secretariat, University of Illinois, 1987.

/4/ TAURIAINEN, A. (ed.), "International Beacon Satellite Symposium .., Proceedings", Oulu (Finland), University of Oulu, Printing Center, 1986.

/5/ SCHMERLING, E. (ed.), "The Physical Basis of the Ionosphere in the Solar-Terrestrial System", (AGARD Conference Proc. No. 295), Paris, AGARD, 1981.

/6/ -, "Large-scale Processes in the Ionospheric-Thermospheric System", Annls. Geophysicae, 5, No. 6, 1987.

/7/ KAISER, T.R. and H.G. MÖLLER (eds.), Proceedings of the Symposium on Upper Atmospheric Winds, Waves and Ionospheric Drifts, J. Atmos. Terr. Phys., 30, 1968, 657-1065.

/8/ RAWER, K. and K. SUCHY, Radio Observations of the Ionosphere, Handbuch Physik, 49/2, 1967, 1-546.

/9/ RAWER, K. (ed.), Results of Ionospheric Drift Observations, Annls. IGY, 33, 1965.

ABO, T., T. TSUDA, Y. TAKASHIMA, R. ITO and S. KATO, Observations of lower ionospheric wind by the Kyoto meteor Radar, J. Geophys. Res., 85, 1980, 177-184.

BELLCHAMBERS, W.H. and W.R. PIGGOTT, Drift observations during the IGY at Halley Bay,/ 9/, 1965, 278-285.

BERTIN, F., J. TESTUD and L. KERSLEY, Medium scale gravity waves in the ionospheric F-region and their possible origin in weather disturbances, Planet. Space Sci., 23, 1975, 493-507.

BIBL, K. and K. RAWER, Travelling disturbances originating in the outer ionosphere, J. Geophys. Res., 64, 1959, 2232-2238.

BIBL, K., W. PFISTER, F.W. REINISCH and C.S. SALES, Velocities of small and medium scale ionospheric irregularities deduced from Doppler and arrival angle measurements, Space Res., XV, 1975, 405-411.

BIBL, K. and F.W. REINISCH, The universal digital ionosonde, Radio Sci., 13, 1978, 519-530.

BRADLEY, P.N. and W. ROSS, Measurement of the direction of arrival of short radio waves reflected at the ionosphere, Proc. Roy. Soc. (London), A 207, 1951, 251-267.

BRIGGS, B.H., The determination of ionospheric drift velocities from three-receiver fading records, Annls. IGY, 3, 1957, 235-249.

BURNSIDE, R.G., F.A. HERRERO, J.W. MERIWETHER jr. and J.C.G. WALKER, Optical observations of thermospheric dynamics at Arecibo, J. Geophys. Res., 86, 1981, 5537-5546.

ESSEX, E.A. and F.H. HIBBERD, Frequency and spatial correlations of fading radio echoes from the ionosphere, /7/, 1968, 1019-1031.

FINELLI, J.A., The phase and group paths of radio waves returned from regions of the ionosphere, J. Atmos. Terr. Phys., 1, 1951, 353-366.

FORBES, J.M., The equatorial electrojet, Rev. Geophys. Space Phys. 19, 1981, 469-504.

FORBES, J.M., Modelling the propagation of atmospheric tides from the lower to the middle and upper atmosphere, Physica Scripta, T 18, 1987, 240-248.

GAVRILOV, N.M., Study of internal gravity waves in the meteor zone, /3/, 1967, 155-166.

GILLE, J.C., General Nature of acoustic-gravity waves, /2/, 1968, 298-321.

GUSEV, V.D., Vest. Moskva Univ., ser. fiz.-mat., 14, No. 6, 1959, 87-99.

GUSEV, V.L. and S.F. MIRKOTAN, Correlation analysis applied to media with three-dimensional characteristics, (Rez. issl. po programme MGG, Ionosfera i meteory), Isv. Akad. Nauk SSSR, No. 4, 7 , 1960.

HANSON, W.B. and R.J. MOFFETT, Ionization transport effects in the equatorial F-region, J. Geophys. Res., 71, 1966, 5559-5572.

HARNISCHMACHER, E., Vergleich der Driftmessungen von Freiburg und Puerto Rico, Kleinheubacher Ber., 11, 1966, 71-78.

HARNISCHMACHER, E., Results of an experiment concerning the fading pattern, /2/, 1968, 277-297.

HEDIN, A.E., N.W. SPENCER and H.G. MAYR, The semidiurnal and terdiurnal tides in the equatorial thermosphere from AE-E measurements, J. Geophys. Res., 85, 1980, 1787-1791.

HINES, C.O., Internal atmospheric gravity waves at ionospheric heights, Can. J. Phys., 38, 1960, 1441-1481.

HINES, C.O., Gravity waves in the presence of wind shears and dissipative processes,/2/, 1968, 356-363.

HUNSUCKER, R.D., Atmospheric gravity waves generated in the high-latitude ionosphere: a review, Rev. Geophys. Space Phys., 20, 1985, 293-315.

HUNSUCKER, R.D. and J.K. HARGREAVES, A study of gravity waves in ionospheric electron content at L=4, /4/, Pt. 1., 1986, 243-275.

JACOBS, K.-G., Windmessungen durch Radio-Beobachtungen von Meteorstreifen und Vergleich mit Fading-Drift-Ergebnissen, Z. Geophysik, 34, 1968, 423-456.

JALONEN, L., Recent studies of the auroral ionosphere using EISCAT, /4/, Pt. 1., 1986, 305-315.

KALCHENKO, P.V., E.L. KASHCHEEV and A.N. OLEINIKOV, Izv. Akad. Nauk (Atmos. & Ocean physics), 21, No.2, 1985, 123-130.

KALCHENKO, P.V., Characteristics of atmospheric disturbances with a quasi-two-day period, /3/, 1987, 112-118.

KAZIMIROVSKY, E.S., E.I. ZHOVTY and M.A. CHERNIGOVSKAYA, A reference model of horizontal drifts in the E- and F-regions, Adv. Space Res., (in print), 1988, 235-236.

RENEGHEA, T.J., M.E. GARDNER and A. PFISTER, Analysis of ionospheric winds and turtulence over Puerto Rico with a correlation method, J. Atmos. Terr. Phys., 27, 1965, 7-30.

KING, J.W., P.A. SMITH, L. ECCLES and H. HELM, Report RRS/I.M.94, Radio Research Station, Slough (U.K.), 1965.

KLOSTERMEYER, J., Gravity waves in the F-region, J. Atmos. Terr. Phys., 31, 1969, 25-45.

KLOSTERMEYER, J., Thermospheric heating by atmospheric gravity waves, J. Atmos. Terr. Phys., 35, 1973, 2267-2275.

KLOSTERMEYER, J., On parametric instabilities of finite-amplitude internal gravity waves, J. Fluid. Mech., 119, 1982, 367-377.

KNUDSEN, W.C., J. FAKKE, K. SPENNER and V. NOVAK, Retarding potential analyzer for the Pioneer-Venus experiment, Space Sci. Instrumentation, 4, 1979, 351-372.

KNUDSEN, W.C., K. SPENNER, K.L. MILLER and V. NOVAK, Transport of ionospheric $O^+$ ions across the Venus terminator and implications, J. Geophys. Res., 85, 1980, 7803-7810.

KNUDSEN, W.C., K. SPENNER and K.L. MILLER, Anti-solar acceleration of ionospheric plasma across the Venus terminator, Geophys. Res. Lett., 8, 1981, 241-244.

KOHL, H. and J.W. KING, Upper atmospheric winds and ionospheric drifts caused by neutral air pressure gradients, Nature, 206, 1965, 699-701.

KRAUTKRAMER, J., Über Wanderungserscheinungen rascher Feldstärke-Schwankungen von Ionosphären-Echos, Arch. Elektr. Übertragung, 4, 1950, 133-138. /original publication as: Leutsche Luftfahrtforschung, FB 1761, T943/.

LIU, C.H., J. KLOSTERMEYER, K.C. YEH, T.E. JONES, T. ROBINSON, G. HOLT, R. LITTINGER, K. SINNO, S. KATO, T. OGAWA, A.J. BEDARD and L. KERSLEY, Global dynamic response of the atmosphere to the eruption of Mount St. Helens on May 18 1980, J. Geophys. Res., 87, 1982, 6281-6290.

MASSEBEUF, M., I. BERNARD, J.L. FELLOUS and M. GLASS, Simultaneous meteor radar observations at Monpazier (France, 44°N) and Punta Borinquen (Puerto Rico, 18°N), J. Atmos. Terr. Phys., 43, 1981, 535-542.

MILGEY, J.E. and H.E. LIEMOHN, Gravity waves in a reslistic atmosphere, J. Geophys. Res. 71, 1966, 3729-3748.

MIES, R.F., The occurrence of parametric instabilities in finite-amplitude internal gravity waves, J. Fluid. Mech., 78, 1976, 763-784.

MILLER, K.L., J.E. SALAH and D.G. TORR, The effect of electric fields on measurements of meridional neutral winds in the thermosphere, /6/, 1987, 337-341.

MOLLER, H.G., Simultaneous observations of meteor winds and ionospheric drifts, /7/, 1968, 701-706.

NAKATA, I., M. KAN and H. UYEDA, Sweep frequency h't measurements of the ionosphere, Rep. Ionosph. Res. Japan, 7, 1953, 129-135.

PFISTER, W. and K. PIBL, A modernized technique for ionospheric drift with spectral analysis, Space Res., XII, 1972, 975-982.

PHILLIPS, G.J. and M. SPENCER, The effects of anisometric amplitude patterns in the measurement of ionospheric drifts, Proc. Phys. Soc. (London), 68B, 1955, 481-492.

ROEVERLEIN, H. and E. NESKE, Ionospheric disturbances of magnetospheric origin, /5/, 1981, 7-1 - 7-11.

SAFARYAGIN, Yu.I., The climatic wind regime in the lower thermosphere from meteor radar observations, /3/, 1987, 31-44.

SANZI, I., Recording wireless echoes at the transmitting station, Nature, 183, 1959, 174.

SAUER, K., Ionospheric irregularities and movements, and: Summary: Irregularities and movements, Proc. Internat. Conference on the Ionosphere, London: The Institute of Physics and the Physical Society, 1963, 239-251 and 370-374.

SAUER, K., Intercomparison of instantaneous drift data, /9/, 1965, 293-296.

SAUER, K. and K. BUCHY, Correlation analysis of an irregular pattern, /8/, 1967, 422-433.

SAUER, K., Ionosphere: Drift measurements, Annls. IGSY, 1, 1968, 117-125.

SAUER, K., Modelling of neutral and ionized atmospheres, Handbuch der Physik, 49/7, 1984, 223-535.

SEINISCH, B.W., New techniques in ground-based ionospheric sounding, Radio Sci., 21, 1986, 331-341.

SEINISCH, B.W., J. FUCHAU, E.J. WEBER, C.G. DOZOIS and K. PIBL, University of Lowell, Center for Atmosph. Res., High Latitude F-Region Drift Studies, AFGL-TR-87-0154, 1986.

SISHBETH, H., L.R. MEGILL and J.H. CAHN, The effect of ion-drag on the neutral air in the ionospheric F-region, Annls. Géophys. 21, 1965, 235-244.

ROBLE, R.G., R.E. DICKINSON and E.C. RIDLEAY, Seasonal and solar cycle variations of the zonal mean circulation in the thermosphere, J. Geophys. Res., 82, 1977, 5493-5504.

ROBLE, R.G., J.E. SALAH and B.A. EMERY, The seasonal variation of the diurnal thermospheric winds over Millstone Hill during solar cycle maximum, J. Atmos. Terr. Phys., 39, 1977, 503-511.

SALAH, J.E., G. HERNANDEZ, R.G. ROBLE and B.A. EMERY, Longitudinal and latitudinal dependences of thermospheric winds from incoherent scatter radar and optical emission Doppler shift measurements, /6/, 1987, 359-362.

SALAH, J.E., M.-L. DUBOIN and Ch. MAZAUDIER, Ionospheric electrodynamics over Saint Santin and Millstone Hill during 26-28 June 1984, /6/, 1987, 351-357.

SCHUNK, R.W., Transport equations for aeronomy, Planet. Space Sci., 23, 1975, 437-485.

SHARADZE, Z.S., G.P. KIKHVILASHVILI, Z.L. LIADZE and N.V. MOSASHVILI, Internal gravity waves in the meteor zone in the Tbilisi region, /3/, 1987, 200-203.

SPIZZICHINO, A., Mesures de vents par météores, /2/, 1968, 201-236.

SPRENGER, K., Die jahreszeitlichen Anomalien der Driftrichtung in der tiefen Ionosphäre nach Beobachtungen auf 245 kHz, Z. Met., 14, 1960, 6-16.

SPRENGER, K. and R. SCHMINDER, On the significance of ionospheric drift measurements in the LF range, /7/, 1968, 693-700.

SPRENGER, K. and R. SCHMINDER, On some relationships between correlation analysis and similar fade analysis results of drift measurements in the lower ionosphere, J. Atmos. Terr. Phys., 31, 1969, 1085-1098.

SUCHY, K., Correlation velocities, Annln. Physik, (7) 18, 1966, 78-96.
UNTIEDT, J., A model of the equatorial electrojet involving meridional currents, J. Geophys. Res., 72, 1967, 5799-5810.
VAN EIKEN, A.P., H. RISHBETH and I.W. WILLIS, Initial EISCAT observations of plasma convection at invariant latitudes 70° - 77°, J. Atmos. Terr. Phys., 46, 1984, 653-641.
VILA P., Intertropical F2 ionization during June and July 1966, Radio Sci., 6, 1971, 689-697.
VOLLAND, H., Electrodynamics of the lower ionosphere, /5/, 1981, 11-1 - 11-13.
WALTEUFEL, P and J.F. McCLURG, Preliminary comparisons of middle and low latitude Thomson scatter data, Annls. Geophys., 25, 1969, 785-793.
WRIGHT, J.W., Interpretation of ionospheric radio drift measurements - 1, /7/, 1968, 919-930.
TOLSTOY, I., The theory of waves in stratified fluids including the effects of gravity and rotation, Rev. Mod. Phys., 35, 1963, 207-230.

Acknowledgement: With kind permission, Figures 2 .. 7, 22, 23 and 28 were copied from "Handbuch der Physik" (vol. 49), Springer-Verlag, Berlin; also Figures 12, 27, 31 .. 33 and 36 from "Winds and Turbulence in Stratosphere, Mesosphere and Ionosphere", North-Holland Publ. Comp., Amsterdam.

# GLOBAL IONOSPHERIC DYNAMICS : A REVIEW

K. Rishbeth

Department of Physics, University, Southampton SO9 5NH United Kingdom

SUMMARY

The ionosphere is a highly dynamic region of the Earth's environment, and its spatial and temporal variations are extremely complex. Practical needs call for a good description of the ionosphere and reliable predictions of its behaviour. This paper reviews some of the scientific problems and the principles behind them.

For practical and scientific purposes, the ionosphere is often "modelled". The most important *parameter for modelling is the electron density distribution* N(h,t) which, *at any one place*, is primarily a function of height (h) and local time (t), but varies also with time of year, magnetic disturbance and solar activity. Thus timescales ranging from minutes up to a solar cycle need to be considered. Many other parameters (such as temperature, composition and velocity) can be modelled and are needed for a physical understanding of the ionosphere.

Aspects of ionospheric structure and behaviour that are considered include the thermosphere wind system and its effects; sporadic E; the low latitude ionosphere; the high latitude ionosphere; and ionospheric storms. The references are illustrative rather than comprehensive.

## 1. INTRODUCTION

An AGARD EPP Symposium is an opportunity to bring together the "scientific" and "practical" aspects of ionospheric radio science and engineering. These aspects originally evolved together, but various factors have tended to drive them apart in the last few decades, and we see increasing specialization, crowded on the engineering side by the explosion of electronic technology, and on the *scientific side by the explosion of knowledge at the time of the IGY, largely through the use of* satellites, rockets and incoherent scatter radar. Ionospheric science - the application of physics to the science of the upper atmosphere - partly became the ionization, being easier to detect experimentally than the neutral gas, and as a "tracer" for the upper atmosphere as a whole, especially the dynamics. The growth of knowledge in aeronomy, especially about the neutral atmosphere, has led to the use of ionospheric theory from the middle 1960's onwards. As the wider field of aeronomy has developed, ionospheric science and aeronomy have taken their place, along with the science of the magnetosphere, the interplanetary medium and the Sun.

Progress in this direction, however, seemed to reduce contact between the "scientific" and "practical" sides of ionospheric science. A contributory factor was the change to the use of communications and the development of adaptive systems that respond to the variable ionosphere. Nevertheless, demands still exist for long-term "predictions" and short-term "forecasts". The problems, traditions, and there is plenty of scope for applying the methods of ionospheric science and fundamental physics to practical needs.

*By convention, the ionosphere* is defined as the part of the upper atmosphere in which free electrons are sufficiently numerous to influence the propagation of radio waves. The terms "region" is used to denote parts of the upper atmosphere (D, 60-90 km; E, 90-140 km; F, above 140 km) and "layer" is used to denote the ionization, e.g. E1 and E2 layers within the E region. For practical purposes the lower limit of the ionosphere can be taken to be at a height of 60 km, and the upper limit where plasma becomes dominant over the neutral gas. Regions of the E2 layer, at 200 km or so, though dense by most standards, are not well-defined, the ionization of a very tenuous constituent of the upper atmosphere (less than 1% of the total density) even at the F2 layer.

To understand the ionosphere, therefore, one must understand the nature of the upper atmosphere in which it exists. All but the lowest part of the ionosphere lies within the thermosphere, the region above 85 km which is strongly heated by solar extreme ultraviolet (EUV) and X-rays, both of which ionize the atmospheric gases. The thermosphere is a huge heat engine driven by solar, auroral and interplanetary sources; the winds - together with associated vertical movements and changes in composition - form a global convective circulation that carries energy away from the heat sources to be radiated elsewhere. Ionospheric motions are largely driven by this engine.

The ions and electrons in the ionosphere are also strongly controlled by the Earth's magnetic field. In low and middle latitudes the geomagnetic field lines are closed, but at higher latitudes they are linked to the magnetosphere. The high latitude ionosphere is profoundly affected by charged particles and electric currents, originating in the magnetosphere and entering the atmosphere mainly in the auroral oval, causing the aurora and magnetic storms. As discussed in other papers, satellite data now begin to follow solar disturbances and the resulting strengthening of the solar wind.

## 2. IONOSPHERE MODELLING AND MODELS

### 2.1 Models of the Ionosphere

The most important ionospheric parameter for the communication engineer is the critical frequency of the plasma, or the electron density N, near the F2 layer peak, the maximum. Modelling is concerned with how N varies with height h, and time t, at a given place, and with latitude and longitude (fig. 1). A starting point in modelling the ionosphere is therefore the creation of a database - an organized collection of data, for example the routine observations of critical frequencies and related parameters at the ionospheric World Data Centres [1], which have been increasingly computerized, and operated variously by the USA and the USSR.

Given a database, the electron density or any other ionospheric parameter can be represented either by purely descriptive "empirical" models, in which mathematical formulas are constructed to match the observed variations of N, or else by "physical" models generated by solving the continuity equation (see Table 1) [2]. An "empirical" or "descriptive" model comprises mathematical formulas or computational algorithms, with empirically determined numerical constants. These models seek to represent the observed values of physical parameters as closely as possible, but are not necessarily based on physics. A very comprehensive example is the International Reference Ionosphere (IRI- [3]. A comprehensive empirical model of the neutral thermosphere is the MSIS global model [4], fitted to worldwide data yielded by Mass Spectrometers aboard satellites and Incoherent Scatter radars on the ground. The model is based on formulas that represent the global temperature distribution (Sec. 6.1). Physical principles, such as the equation of state and the barometric law, are then used to compute gas densities and other parameters in the MSIS model.

A "physical" model is theoretical or computational in nature. It simulates the behaviour of the ionosphere, ie electron density by solving the physical equations that govern it: the equations of state, continuity, motion, energy. Apart from the task of setting up and solving the equations, the main problems are to include all the important processes; obtain good values of the various parameters; and to set the right boundary conditions. The more sophisticated models solve equations, not only for the density N but also for the velocity V and temperature T of the electrons and ions (Table 1). The comprehensive "coupled" models treat also the corresponding equations for the neutral air, in order to take account of its important interactions with the electrons and ions. The development of physical models of the ionosphere, "coupled" and otherwise, is a major activity of ionospheric modelling. Many ionospheric models of such types have been developed by U.S. and European groups, e.g. [5,6]. A useful compromise between the "empirical" and "physical" approaches is represented by "semi-empirical" models, which simplify the equations to make them computationally fast to solve, and are then used to generate the numerical coefficients for an empirical model, e.g. [7,8].

TABLE 1 : BASIC EQUATIONS FOR ANY KIND OF PARTICLE

CONTINUITY EQUATION (conservation of mass)

[RATE OF CHANGE OF ELECTRON DENSITY] [PRODUCTION] - [LOSS] - [TRANSPORT]

EQUATION OF MOTION (conservation of momentum)

[ACCELERATION] [DRIVING FORCE] - [DRAG] - [ADVECTION]

HEAT BALANCE EQUATION (conservation of energy)

[RATE OF TEMPERATURE CHANGE] [HEATING] - [COOLING] - [CONDUCTION]

ADVECTION: transport of momentum by viscosity, etc

## 2.2 Regular and Irregular Variations

The real ionosphere has often been discussed in terms of the simple Chapman layer model [9], in which the electrons recombine by a "square law" $\alpha N^2$, and the critical frequency depends on solar zenith angle $\chi$ and mean sunspot number R according to the formula

$$f_0 = A\,(1 + aR)\,(\cos\chi)^n \qquad [n = 0.25] \qquad (1)$$

where A is a numerical coefficient. Besides being the first real "physical" model of the ionosphere, (1) is used "empirically" by fitting A and n, as parameters, to actual data. As is well known, the daytime E and E1 layers follow the formula (1) quite well, whether $\chi$ varies with time of day, season, or latitude. The index n varies between about 0.2 and 0.3, departures from n = 0.25 being attributable, physically, to the variation of temperature with height, and to perturbations due to drift and other factors. The incident flux of solar ionizing radiation varies with the solar cycle, and so therefore does $f_0$. Fig. 2 shows the regular seasonal and solar cycle variations of the E and E1 layers.

Like most natural systems, the ionosphere shows a great deal of variability even under magnetically quiet conditions. The normal E layer is the least variable, being reasonably predictable on an hourly and daily basis at a set for a given level of solar activity. Variability is greatest for the E1 layer, being more than 10% in $foE^2$ from day to day, with plasma-physics variations that have scale-size features. By way of illustration, Fig. 3 shows the variation of $foE^2$ throughout complete years of unspot min ma and sunspot maxima; the seasonal and day effects are obvious, and many other features can be found.

## 3. BASIC PHYSICAL PROCESSES OF THE IONOSPHERE

As a prelude to discussing some major problems and anomalies of the ionosphere, this section summarizes the theoretical background. Where no specific citations are given, the material may be found in standard texts, such as those listed in the Bibliography.

### 3.1 The Continuity Equation

The continuity equation (Table 1) for electrons and ions is

$$\partial N/\partial t = q - L(N) - \mathrm{div}(NV) \qquad (2)$$

Many useful deductions can be made by considering this equation and its terms in general, without solving it in detail [10], for example:

(a) Since many of the physical quantities involved depend quite critically on atmospheric pressure or density, it is often simpler, and physically meaningful, to solve (2) in terms of fixed pressure-levels in the atmosphere, rather than fixed heights.

(b) Major perturbations of the F2 layer generally come about through changes in the production and loss terms. Vertical drift is effective mainly because it moves ions and electrons from where they were formed, to levels where the production/loss balance is different.

(c) Horizontal motions are not usually very important in the continuity equation, *except* in special circumstances such as (i) near the magnetic equator, or (ii) where there already exist large horizontal gradients of electron density (e.g. because of particle precipitation).

(d) In consequence, the transport term in (2) may usually be simplified by omitting horizontal motion, and retaining only the vertical drift velocity $w$, in which case this term reduces to

$$\mathrm{div}(NV) = \partial(Nw)/\partial h \qquad (3)$$

### 3.2 Production and Loss

Although the subject may not seem within the scope of this paper's title, a summary of some basic facts is appropriate. Detailed numerical data may be found in references [11-13], for example. The ordinary ionospheric layers are produced by extreme ultraviolet (EUV) and X-rays, emitted both as spectral lines and continuum radiation by the Sun. The principal radiations that produce the layers are: E1 layer (150-180 km): EUV 17-80 nm (the softest e.m. radiations); E layer (100-120 km): X 1-17 nm & EUV 81-103 nm (Ly $\beta$ 102.6 nm); D layer (70-90 km): Ly $\alpha$ 121.6 nm & X-rays 0.1-1 nm; $\zeta$ layer (50-70 km): MeV cosmic rays (not a significant layer).

The heights are approximately those where the radiations are most strongly absorbed. There exists no radiation that is strongly absorbed above the E1 layer; the E1 and F2 layers are formed by the same radiation, the greater electron density of the F2 layer being due to a smaller rate of loss of electrons. The loss term $L(N)$ in (2) can be approximated thus:

$$\text{E layer...}\quad L(N) = \alpha N^2 \qquad \text{F2 layer...}\quad L(N) = \beta N \qquad (4)$$

where $\alpha$, $\beta$ are the square-law and linear loss coefficients. For the E and E1 layers, the effective square-law loss coefficient is a combination of the values of $\alpha$ for the molecular ions present. For the F2 layer, the linear loss coefficient is

$$\beta = k' n[O_2^+] + k'' n[N_2] \qquad (5)$$

where $k'$, $k''$ are the rate coefficients of the relevant chemical ion-atom interchange reactions. Whether or not the E1 layer appears as a distinct layer can be shown to depend on the ratio $G = \beta^2/\alpha q$, evaluated at the level of peak production in the E1 layer.

### 3.3 The F2 Peak

The upward increase of electron density, above the E1 layer, stops at the F2 peak, the height where chemical control gives way to gravitational or diffusive control. At the F2 peak, the transport terms in the continuity equation (2) are comparable to production and loss, which are roughly in balance below that height. At greater heights, ions and electrons take up a gravitational distribution, e.g. [14]. The F2 peak is governed by the relations (in which D is the ion-neutral diffusion coefficient, and the suffix $_2$ denotes values at the peak height $h_m F2$):

$$D_2 = D_1 H^2 \qquad N_2 = q_m H_2 \qquad (6)$$

The height of the peak can be altered by a neutral air wind or an electric field. A horizontal wind blowing towards the magnetic equator drives the ionization up magnetic field lines, raising the peak and increasing $N_m$. Opposite effects are produced by a poleward wind. Poleward winds tend to occur at night (at least for the equinox situation of Fig. 4), equatorward winds by day. In principle, eastward/westward electric fields also produce upward/downward electromagnetic drift.

Some of the most challenging problems are in the field of F2 layer behaviour. Physically, this is studied by solving the continuity equation (2), in which V is the transport velocity due to winds, electric fields and plasma diffusion. As is well known, the solutions are complicated to obtain, and difficult to "calibrate" in terms of real data. Perhaps the main reason is that the physical parameters and coefficients in the equation are not easily measured, and most are known only to within an accuracy

of say 25%. Production, loss and diffusion coefficients of course depend on the neutral air density, so any useful physical solution of (2) requires such knowledge, which can only come from a reliable "empirical" or "physical" model of the neutral upper atmosphere.

For mid-latitudes, the "servo" model [15] represents the layer as a quasi-equilibrium system in which the height ($h_m$), responding to vertical drift, tends at any time towards an "equilibrium" value ($h_e$), primarily determined by diffusion and loss rates, but is modified by vertical drift W(t), in the manner represented by the "hunting equation"

$$dh_m/dt = D'(h_e - h_m) + W(t) \qquad (7)$$

(where D' is related to the diffusion coefficient) and the time variation ($dN_mF2/dt$) of the peak density is then computed from the ambient production and loss rates at height $h_m$. The "servo" equations that describe this behaviour are much simpler to solve than the full continuity and, if properly calibrated, might in principle provide a useful tool for F2 layer propagation calculations.

### 3.4 Topside Ionosphere

The shape of the topside F2 layer is controlled by diffusion. The topside is not static, as there is generally some flow of plasma along geomagnetic field lines. This is important for the maintenance of the F2 layer at night: within limitations [16], the F2 layer supplies ions, through $O^+ - H^+$ charge exchange, to the overlying protonosphere by day, with a reverse downward flow at night. The protonosphere may thus act, to a limited extent, as a reservoir. In high latitudes, the ionosphere can be an important source of magnetospheric ions [17].

## 4. UPPER ATMOSPHERIC INFLUENCES ON THE IONOSPHERE

### 4.1 Thermospheric Structure and Composition

Up to about 100 km the major gases (molecular nitrogen and oxygen) are well mixed by winds and turbulence, and their concentration ratio does not change with height, though this is not the case for the many minor constituents. Turbulence ceases at the turbopause near 100 km, and above 100 km the gases are diffusively separated by the action of gravity, so that atomic oxygen (formed by the dissociation of oxygen molecules by solar UV radiation) progressively becomes more abundant than the heavier molecular gases. The chemical composition also varies with time and season. For example, the atomic molecular ratio above 200 km is greater in winter than in summer, while localized variations of composition occur in auroral latitudes. These variations arise in a complex way from the global thermospheric circulation driven by solar and auroral heating. The modelling of thermospheric dynamics and composition has progressed far in recent years, e.g. [15-21].

The temperature of the thermosphere is governed, at the bottom, by the mesopause temperature (about 180 K), but it increases rapidly upwards with a gradient dT/dh determined by the solar heat input. Higher up, where the thermal conductivity of the air becomes large, the temperature gradient dT/dh flattens out, and T tends to a limit known as the exospheric temperature $T_\infty$, which varies considerably with local time (LT), latitude, season and the 11-year sunspot cycle. Typically at mid-latitudes $T_\infty$ is 800 K at midday and 600 K at midnight at sunspot minimum, and 1400 K and 1100 K respectively at sunspot maximum.

The daily maxima and minima of $T_\infty$ (around 1400 and 0100 local time) are situated in low latitudes near equinox but migrate north and south at the solstices (Figs. 3 and 4). The temperature variations are mainly due to solar heating, which causes the atmosphere to expand or bulge on the daylit side of the Earth, creating a high-pressure region often referred to as the daytime pressure bulge. Localized heating in the auroral ovals, due to the energy deposited by energetic particles and electric currents from the magnetosphere, is not shown in these maps.

### 4.2 Upper Atmosphere Winds

The temperature variations shown in Figs. 3 and 4 give rise to corresponding pressure gradients. Just as in the lower atmosphere, horizontal variations of air pressure produce winds. The wind velocity U of the neutral air is controlled by the Coriolis force due to the Earth's rotation with angular velocity $\Omega$; the molecular viscosity of the air (coefficient $\mu$); and ion-drag - the friction due to collisions between air molecules and ions, which exists because the ions cannot move freely with the wind. The ion-drag depends on the ion-neutral collisions rate coefficient $k_i$ and the difference between the wind velocity U and the ion drift velocity V, the latter being mainly controlled by electromagnetic forces (Sec. 4.5). Gravity and the vertical pressure gradient are omitted because they are almost precisely in balance. The equation for the horizontal wind velocity is then

$$dU/dt = F - 2\,\Omega \times U - \mu \nabla^2 U - \rho N_i k_i (V - U) \qquad (8)$$

where

$$F = -\nabla_{horiz} p / \rho \qquad (9)$$

represents the effect of the horizontal pressure gradient. At great heights the viscosity becomes very effective and tends to destroy spatial variations of U; as a result, the wind [illegible] little with height. Lower down, small-scale velocity gradients are not smoothed out by viscous

The wind direction depends on the ratio of Coriolis force to ion-drag, as may be illustrated by considering some special steady-state cases (with $dU/dt = 0$ and viscosity neglected). If Coriolis force is dominant, and ion-drag is small as in the lower ionosphere, the wind blows at right angles to the pressure gradient. Then, at latitude $\phi$:

$$U = F/(2\Omega \sin\phi) \quad (\underline{U} \perp \underline{F}) \qquad (10)$$

(the situation familiar in the weather maps of the lower atmosphere, with the wind blowing along the isobars). A different situation exists in the daytime F layer, where ion-drag is large and $\underline{U}$ is almost parallel to $\underline{F}$:

$$U = F/(KN \sin I) \quad (\underline{U} \parallel \underline{F}) \qquad (11)$$

If both ion-drag and Coriolis force are significant, the wind is inclined to the force F. Accordingly, the schematic wind vectors shown in Figs.4 & 5 are nearly parallel to the temperature (and pressure) gradient by day, but are slightly deflected by Coriolis force at night.

The neutral air velocity is of course subject to equations of continuity and energy, as well as the equation of motion (8). Generally speaking, production and loss processes are unimportant for the major constituents of the neutral air, and the continuity equation (2) is just

$$\partial n/\partial t = -\operatorname{div}(n\underline{U}) \qquad (12)$$

The pressure distribution and the wind velocity automatically adjust themselves in such a way as to satisfy this equation. Thus any divergence (or convergence) of the horizontal winds must be balanced by upward (or downward) winds, so that div ($n\underline{U}$) is small. Furthermore, the motion must satisfy the energy equation.

## 4.3 Atmospheric Tides

The thermospheric winds just described are part of a very complex system of atmospheric tides, which can be resolved into numerous components (24-hour, 12-hour, etc with different types of latitude variation) [22]. The tides are "forced" (driven) by the heating effect of the Sun, and by the gravity of the Sun and Moon, though (unlike the case of marine tides) the gravitational tides are very much weaker than the predominant solar thermal tide, which is driven by:

--- Absorption of solar EUV and X-rays in the thermosphere;
--- Absorption of solar UV in the ozone layer;
--- Heating of the ground and lower atmosphere by visible and infra-red radiation.

The weak lunar tides in the thermosphere gives weak ionospheric and geomagnetic effects. The most important effect of tides on the ionosphere is the generation of electric fields and current by dynamo action in the geomagnetic field, giving rise to electromagnetic drifts (Sec. 4.5).

## 4.4 Gravity Waves

The tides may be thought of as a special kind of atmospheric wave, of which the atmosphere contains a vast array. The wave periods range over several orders of magnitude, from the 26-month quasi-biennial oscillation to infrasonic waves of periods <1 s (some of which are associated with aurora and with seismic activity). Important types in the upper atmosphere include planetary waves, with wavelengths of order 1000km and periods of days, which are important in the mesosphere; tides (Sec. 4.3); and acoustic and gravity waves. Acoustic waves are oscillations of relatively short period (< 4 min), in which the restoring force is due to compression and gravity is unimportant. Gravity waves are oscillations controlled by the buoyancy of the air and have a complicated phase structure. They can be divided into classes:

--- Short-period and medium-period gravity waves (periods of 5-30 minutes, wavelengths of 100-300 km, horizontal speeds of 100-200 ms-1) are generated in the lower atmosphere, for example by storms, by winds blowing over mountains, and occasionally by man-made explosions.

--- Long-period gravity waves (periods of 0.5-3 hours, horizontal wavelengths of 1000-4000 km, speeds of 400-700 m s-1) are generated by disturbances in the thermosphere, mainly in the auroral oval. They can travel for thousands of kilometres, and are particularly prevalent at times of geomagnetic disturbance.

Waves propagated to the thermosphere from lower heights are selectively filtered by the temperature structure of the mesosphere, in a way that varies with season. Herein lies a strong source of variability.

## 4.5 Electric Fields and Drifts

If the ions are drifting under the influence of an electric field, generated by dynamo action or introduced from the magnetosphere, the ion-drag term in (8) acts as a driving force. Similar considerations to those of Sec. 4.2 apply to the resulting wind direction. In high latitudes, these "electrically-driven winds" may predominate, though auroral heating produces localized pressure gradients and associated winds, so both causes of motion must be fully considered in the dynamics.

Except at low latitudes, the "electromagnetic drift" $\underline{E} \times \underline{B}/B^2$ that is produced by an electrostatic field $\underline{E}$ in the F region, is predominantly horizontal and therefore does not contribute much to the continuity equation (2). It can however have important consequences in the large-scale "convection" pattern at high latitudes (Sec. 5.4), because of the way the convective drift carries ions and electrons between regions of very diverse production and loss rates. At low latitudes, the "electric fountain" that produces the "F2 layer equatorial anomaly" is a striking example of electric field effects. Electric fields - whether tidal or magnetospheric in origin - are very variable, and probably contribute greatly to the day-to-day variability of the F layer.

## 5. *IONOSPHERIC PHENOMENA*

Although the behaviour of the mid-latitude quiet ionosphere is quite well known and understood, many phenomena that bear on radio propagation are not fully explained, and there are many problems at high and low latitudes. Ionospheric variability (hour-to-hour, day-to-day or place-to-place) is a pervasive problem, and many of the topics to be discussed have some bearing on it. Only a selection of topics that are particularly associated with dynamics are mentioned here and (as before) "textbook" material can be found in works listed in the Bibliography.

### 5.1 Sporadic E

Sporadic E (Es) is important to radio propagation. High Es critical frequencies, occurring in the sporadic and unpredictable manner that gives sporadic E its name, can present severe problems to radio communications. It is thus worth considering what physical understanding exists of Es, in the hope that understanding might eventually lead to some degree of predictability.

Twelve standard classifications of sporadic E, based on the appearance of ionograms, have been defined [23]. These classifications are morphological rather than physical, and their usefulness is be limited. Physically there may be only three main types: (1) Equatorial Es, a plasma instability caused by the high electron drift velocity (i.e. the large current density) in the diurnal equatorial electrojet. Its physics seems reasonably well understood [24] though interesting questions remain, especially regarding the "counter-electrojet" which is associated with some other ionospheric phenomena. Since it is fairly regular and its properties are well known, the equatorial electrojet does not present any great practical problem.

(2) Particle Es, caused by precipitation of auroral particles, which has serious effects on propagation at high latitudes, but might be treated as an aspect of the auroral oval and its associated phenomena.

(3) Mid-latitude Es is thought to be produced by wind shears, i.e. small-scale gradients of wind velocity [25]. The sheared wind interacts with the geomagnetic field in such a way as to compress ions into thin layers, typically 1 km thick but 100 km in horizontal extent. Theory shows that, to exist for minutes or hours as observed, these layers must comprise long-lived ions rather than the ordinary gaseous ionospheric ions. It is generally accepted that these ions are metallic, formed by ionization of trace constituents such as Fe, Mg and Ca. Some forms of these layers are associated with semidiurnal tidal winds, which are fairly regular and might have some predictability, others are forced by small-scale winds, possibly by gravity waves, and for practical purposes are truly "sporadic". So, even where the likelihood of Es is statistically known, detailed predictions seem beyond reach at present.

### 5.2 Mid-Latitude F2 Layer Anomalies

The F2 layer is not very regular, and in Figs. 1 - 3 has some features that differ from those of a purely solar-controlled layer, the main anomalies, with their probable causes, are as follows:

(a) The shape of the day/night variations of $N_mF2$ and $h_mF2$ may largely be due to neutral air winds, and (to some extent) electric fields. Around sunset, the meridional component of the wind changes from poleward to equatorward - see Fig. 4 - and the resulting F2 layer drift changes from downward to upward. This is a probable cause of the commonly observed increase of $N_mF2$ in the late evening. The often pre-dawn decrease of $N_mF2$ is probably connected with the equatorward-to-poleward wind reversal around that time.

(b) The anomalous seasonal variation (Fig. 1) is attributable to changes of the atomic/molecular ratio of the neutral air, produced by the global circulation in the thermosphere.

(c) The semiannual variation in the F2 layer is a manifestation of a global thermospheric phenomenon, the origin of which may lie in interactions with lower levels in the atmosphere.

(d) The F2 layer is maintained at night, partly by being raised by winds (which reduces the effective value of loss coefficient in (5)), partly by inflow from the protonosphere, and partly by ionization by nocturnal EUV sources and perhaps particle precipitation.

### 5.3 The Equatorial Ionosphere

*In low latitudes, electric fields* cause a drift of ionization away from the magnetic equator by day, producing a depletion known as the equatorial trough or Appleton anomaly. Intense irregularities, known as equatorial spread F, may occur in the equatorial F layer, generally after sunset when plasma instabilities develop. The other major phenomenon is the equatorial electrojet in the E layer, which depends on strong polarization fields that develop in the equatorial situation. Current problems are concerned with day-to-day and solar cycle variabilities, the interaction between the ions and neutral atmosphere, and the mechanisms of penetration of magnetospheric electric fields to low latitudes [26].

### 5.4 Auroral Oval and Associated Phenomena

Many geophysical phenomena are caused by auroral particle precipitation, the incidence of which is closely related to the "auroral ovals", rings 2000-3000 km in diameter that surround the magnetic poles. Two types of precipitation have been defined [27], the intense impulsive "splashes" of relatively soft (keV) particles associated with substorms, and the more persistent background "drizzle" of harder (MeV) particles at rather lower latitudes. They include:

| SPLASH EVENTS | DRIZZLE EVENTS |
|---|---|
| Discrete auroral forms | Steady diffuse aurora |
| Auroral absorption events | Slowly-varying absorption |
| Auroral sporadic E | Low diffuse sporadic E |
| Rapidly fading VHF scatter | VHF forward scatter @ 85 km |
| Impulsive Pi pulsations | Continuous Pc pulsations |
| Bursts of VLF auroral hiss | Quasi-constant VLF chorus |
| Short bursts of keV electrons | Sustained electron fluxes |
| Bremsstrahlung X ray bursts | Long-duration hard X rays |
| Spread F echoes | |
| Negative magnetic bays | |

Since many of these phenomena influence radio wave propagation, any scheme that organizes the data on their occurrence is potentially useful. In principle it should be possible to represent the *position of the auroral oval in terms of geographic, geomagnetic and solar-terrestrial parameters, either* "empirically" or "physically" (in the sense explained in Sec. 2.1). Then the positions of ionospheric features could be related to the geometry of the auroral oval.

Fig. 6 is a schematic plan of the northern auroral oval and features linked to it. The oval can in fact be quite well represented by a circular ring, the centre of which is offset by a few degrees on the midnight side, from the magnetic dip pole. Both the mean latitude (M) and offset (S) depend on the geomagnetic K-figure, which in turn depends on solar activity and the configuration of the interplanetary magnetic field. This leads to the following equations for the geomagnetic latitude of the oval, in terms of local geomagnetic time GT (in hours), of the type

$$L = M - S \cos \pi (GT)/12 \qquad (1)$$

Suggested values (in degrees) for the boundaries of the oval are:

| | | | | |
|---|---|---|---|---|
| Poleward boundary: | M | $73.7 - 0.1\ K$ | S | $3.0 + 0.6\ K$ (1a) |
| Equatorward boundary: | M | $71.5 - 0.5\ K$ | S | $7.6 + 0.15\ K$ (1b) |

The parameters involved (magnetic latitude, magnetic activity and time) can be defined in different ways. If the K-figure is taken as the high latitude quarter-hourly index $K_H$, the equations fit quite well the northern hemisphere auroral data [25], though for other purposes different parameters might fit better.

### 5.3 The Convection Pattern and its Effects

The magnetic linkage between the solar wind and the high-latitude geomagnetic field leads to a fast day-to-night drift of ionospheric plasma across the polar cap [20]. This supplies plasma to the nightside of the polar ionosphere, although this supply may be finely balanced, in that the rapid drift *also leads to an increase in the decay rate of the plasma.* This may solve the old puzzle as to how the ionosphere was maintained throughout the polar winter in the absence of solar ionizing radiation. (The polar caps are not locations of strong particle precipitation, and only weak sources of ionizing radiation exist in the night sky).

This day-to-night drift (sketched in Fig. 6) is part of the large-scale circulation or convection of plasma in closed loops. At sub-auroral latitudes, the convection pattern also seems to account for the trough, a region of low plasma density. In particular, the combination of the Earth's rotation and the convection velocity causes plasma to become almost stationary with respect to the Sun, in the evening sector. If this stagnation point (at 20 GT in Fig. 6) is in darkness, the plasma density can decay to very low levels [10].

## 6. STORMS AND THEIR IONOSPHERIC EFFECTS

### 6.1 Summary of Storm Phenomena

Energetic plasma streams travel to the Earth as strengthenings of the solar wind. Their complex terrestrial effects include:

1. Magnetospheric compression; enhanced energy input leading to frequent and violent auroral substorms; the storm-time ring current.

2. Intensified activity in the auroral ovals, with expansion towards lower latitudes.

3. Geomagnetic disturbances; the storm-time decrease due to the magnetospheric "ring current"; magnetic bays due to auroral currents.

4. *Ionospheric storm effects at all latitudes; short-term and long-term D layer effects;* F layer negative and positive storms.

### 6.2 Phenomena of Ionospheric Storms

The ionospheric effects are widespread, unpredictable in detail, and are most important from the point of view of radio propagation. Some examples are clearly visible in Fig. 2. However, some patterns of behaviour are known, and might in principle provide some predictability, if only on an empirical basis. To some extent the storm phenomena are believed to be transmitted from auroral latitudes towards lower latitudes [31]. Just how much that occurs, and whether it could form the basis of short-term forecasting, is still a matter for investigation.

The F2 layer is often profoundly affected during magnetic storms, with serious consequences for radio propagation. At mid-latitudes the F2 layer electron density increases during the "initial phase", then often shows severe decreases in the "main phase" of the storm. The changes have complex local time and universal time variations. Recovery takes 2-3 days.

In the lower ionosphere, particle precipitation at high latitudes produces changes in the D layer, which are propagated to the mid-latitude D layer, and which may last for ten days or so [32]. The mechanisms of this "storm after-effect" involve atmospheric chemistry and dynamics.

6.3 F2 Layer Storm Physics

The F2 layer storm effects have been attributed to combinations of enhanced thermospheric winds, chemical composition changes, and particle precipitation, all of which are known to occur, and which might in principle result from disturbances propagated from the auroral ovals. In effect the high-latitude energy inputs modify the global thermospheric circulation, in the manner portrayed schematically by Fig. 7 [33-35]. It is well established that the global circulation is severely modified during magnetic storms, and that the world-wide changes of thermospheric temperature are a consequence of that; the "storm circulation" is the global mechanism through which the storm energy is dissipated throughout the atmosphere.

Current work suggests that (a) the best explanation of the "main phase" F2 layer effects lies in increases of the loss coefficient in the continuity equation (2); (b) that these increases can reasonably be attributed to changes of chemical composition, such as are observed; (c) it is possible, though unlikely, that the F2 layer loss coefficient might instead be enhanced by increases of the rate coefficients in (5), perhaps through molecular excitation; (d) there are difficulties in explaining how the "storm circulation" does change the chemical composition in this way; (e) the roles of turbulence and of minor constituents in the lower thermosphere may be very important in controlling the "storm circulation" and its effects; (f) the "positive" main phase storms can be explained in terms of composition changes produced by the storm circulation; (g) energy input - probably particle precipitation - may take place during storms at places well equatorward of the auroral oval, and cause storm effects locally; (h) the possible association between F2 layer depletions and stable red arcs needs more study; (i) the initial "positive phase" may be due to equatorward winds, or perhaps to electric fields [36-38].

A complete physical description of ionospheric storms is still remote, even in a purely theoretical way. Perhaps the ideal to aim for is a theoretical ionospheric model, well calibrated against real data, with appropriate modelling of ionosphere/solar wind/IMF links; and coupled with IMF and solar wind monitoring, and solar flare predictions. Whether that is realizable is conjectural, but much can be learned from attempting it.

7. CONCLUSION

Several questions remain. One concerns the prediction of median quiet-day parameters of the ionosphere, which are useful for long-term planning of services. Since most quiet-day ionospheric phenomena are quite well documented (even if not fully explained) such predictions are largely a matter of predicting solar activity, and choosing the best parameters to relate solar activity to ionospheric parameters such as foF2 [39]. A more difficult question concerns the short-term prediction or "forecasting" of deviations from mean behaviour. In theory, this might be unnecessary if all communications systems were adaptive, and capable of responding in real time to changing conditions but even if this were so, there would still be value in forecasting the likelihood of ionospheric disturbance, which again requires the monitoring of solar and interplanetary conditions. There might still be an operational need for forecasting the progress of ionospheric storms, once they have started; a challenging problem.

To a large extent, the problem lies not in the ionosphere, but well outside it. If a way could be found to link ionospheric model parameters to observable solar or interplanetary quantities (with any necessary time delay), this might form the basis of a forecasting scheme for high latitude propagation characteristics. Of interest and practical value is the fact that some effects of the IMF on the geomagnetic field can be monitored by ground-based magnetometers [the Svalgaard-Mansurov effect].

Many lines of study suggest themselves for the future. For the present, it is vital for solar-terrestrial monitoring and prediction services to be kept in being, if progress in understanding the ionosphere and its variability is to be maintained.

To end by summarizing some major questions for further research:

(1) What drives the global circulation in the upper atmosphere? i.e., all the energy sources and their relative importance [Solar EUV, particles and currents from the magnetosphere and solar wind; energy propagated from below; others?]

(2) Is the ionosphere influenced by the Earth's features (oceans, mountains), or by earthquakes, thunderstorms? If so, how?

(3) Links with the magnetosphere and the solar wind?

(4) Mechanisms of ionospheric storms?

(5) Throughout the ionosphere, there is irregular structure, on scales of a few metres to 100 km. What causes it all?

(6) Active experiments: the ionosphere as a physics/chemistry lab?

(7) Is the ionosphere predictable on time-scales of practical interest?

ACKNOWLEDGEMENTS

Acknowledgement for the diagrams is made to Rutherford Appleton Laboratory, U.K. [Figs. 1,2,4,5,6]; Max-Planck-Institut fuer Aeronomie, Katlenburg-Lindau, F.R.G. [Fig. 3]; and the Americal Geophysical Union, Washington, D.C., U.S.A. [Fig. 7]. Some passages of the text are derived from material in Chapter 5 of "Radio Wave Propagation" (editors M.P.M.Hall & L.W.Barclay), IEE/Peter Peregrinus, Stevenage, U.K., 1985, and in Chapter ~ of "Modern Radio Science" (editor A.L.Cullen),ICSU Press, Paris, France,1987.

BIBLIOGRAPHY

Air Force Geophysics Laboratory (A.S.Jursa, editor), "Handbook of Geophysics and the Space Environment", 1985
P R Banks & G Kockarts, "Aeronomy" (2 vols), Academic Press, New York, 1973
T Beer, "Atmospheric Waves", Adam Hilger, London, 1974.
A Giraud & M Petit, "Ionospheric Techniques and Phenomena", Reidel Dordrecht, 1978.
J K Hargreaves, "The Upper Atmosphere and Solar-Terrestrial Relations", Van Nostrand-Reinhold, New York, 1979
J S Nisbet, "Atmospheres of Earth and the Planets", Reidel, Dordrecht, 245-258, 1974.
J A Ratcliffe, "Sun, Earth & Radio", World University Library, 1970
H Rishbeth & O K Garriott, "Introduction to Ionospheric Physics", Academic Press, New York, 1969
R C Whitten & I G Poppoff, "Introduction to Aeronomy", Wiley, New York, 1971

REFERENCES

1. The World Data Center System (1987) WDC-A, Boulder, Colo.
2. W Kohnlein, Rev. Geophys. Space Phys. 16 (1978) 341.
3. K Rawer, "IRI-79" (1981), Report UAG-82, WDC-A, Boulder, Colo.
4. A E Hedin, J. Geophys. Res. 88 (1983) 10170.
5. T J Fuller-Rowell et al., J. Geophys. Res. 91 (1986) 7711.
6. J J Sojka et al., Planet. Space Sci. 33 (1985) 1375.
7. C M Rush et al., Radio Sci. 19 (1984) 1083.
8. D N Anderson, M Mendillo & B Herniter (1985) AFGRL Report IRI-025.
9. S Chapman, Proc. Phys. Soc. 43 (1931) 26.
10. H Rishbeth, J. Atmos. Terr. Phys. 46 (1984) 511.
11. L Heroux & H E Hinteregger, J. Geophys. Res. 83 (1978) 5305.
12. H E Hinteregger, Adv. Space Res. 1 (1981) 39.
13. D G Torr & M R Torr, J. Atmos. Terr. Phys. 41 (1979) 797.
14. H Rishbeth, Geophys. J. Roy. Astro. Soc. 41 (1975) 311.
15. H Rishbeth, J. Atmos. Terr. Phys. 29 (1967) 225.
16. G J Bailey, R J Moffett & J A Murphy, Planet. Space Sci.26 (1978) 743
17. T E Moore, Rev. Geophys. Space Phys. 22 (1984) 264.
18. T J Fuller-Rowell & D Rees, J. Atmos. Sci. 37 (1980) 2545.
19. R G Roble, R E Dickinson & E C Ridley, J. Geophys.Res. 86 (1981) 1499
20. T J Fuller-Rowell & D Rees, Planet. Space Sci. 31 (1983) 1209.
21. R G Roble & I F Kasting, J. Geophys. Res. 89 (1984) 1711.
22. J M Forbes, Phys. Scr. T18 (1987) 240.
23. W R Piggott & K Rawer (1972), Report UAG-23, WDC-A, Boulder, Colo.
24. D T Farley, J. Atmos. Terr. Phys. 47 (1985) 779.
25. J D Whitehead, Rev. Geophys. Space Phys. 8 (1970)
26. R Raghavarao et al., WITS Handbook, SCOSTEP (1987)
27. T R Hartz & N M Brice, Planet. Space Sci. 15 (1967) 301.
28. Y I Feldstein & G V Starkov, Planet. Space Sci. 15 (1967) 209.
29. W C Knudsen, J. Geophys. Res. 79 (1974) 1046.
30. J J Sojka et al., J. Geophys. Res. 87 (1982) 187.
31. K Davies, Planet.Space Sci. 22 (1974) 237.
32. J S Belrose & L Thomas, J. Atmos. Terr. Phys. 30 (1968) 1707.
33. R A Duncan, J. Atmos. Terr. Phys. 31 (1969) 59.
34. H G Mayr & H Volland, Planet. Space Sci. 20 (1972) 379.
35. H Rishbeth, Radio Sci. 9 (1974) 153.
36. A S Rodger, J. Atmos. Terr. Phys. 46 (1984) 335.
37. G W Prolss, Planet. Space Sci. 35 (1987) 807.
38. H Rishbeth, T J Fuller-Rowell & A S Rodger, Phys. Scr. 36 (1987) 327.
39. P A Smith, J. Atmos. Terr. Phys. 30 (1968) 177.

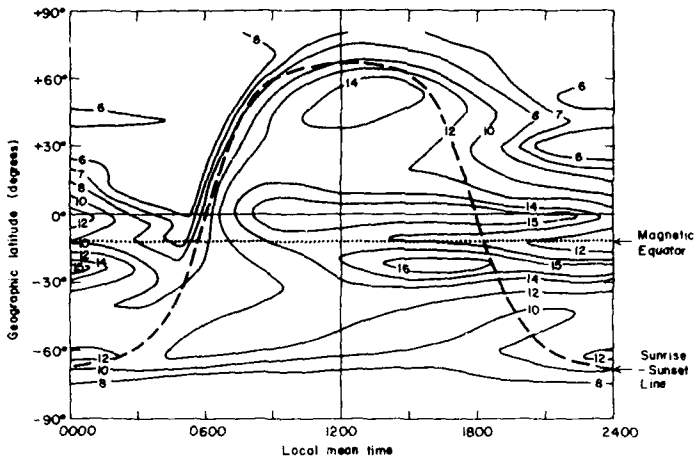



Fig. 1. Prediction map for F2 layer critical frequency, in local time for the American sector, for [illegible], Central Radio Propagation Laboratory, U.S. Department of Commerce.

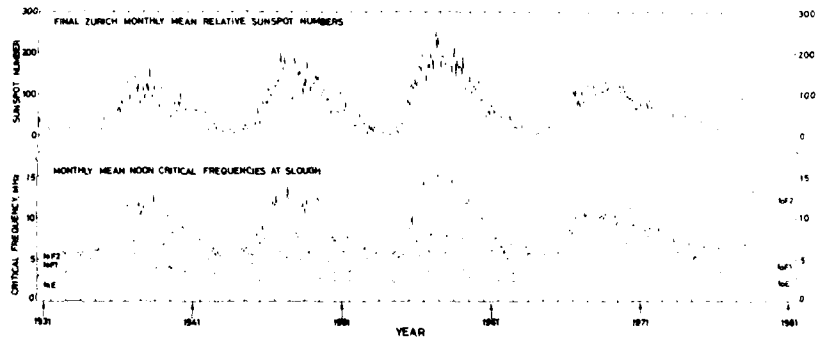



Fig. 2. Noon monthly mean critical frequencies of the F2, F1 and E layers at Slough, U.K. (Lat. 52N°), 1931-1951. Copyright: S.E.R.C. Rutherford Appleton Laboratory, U.K.

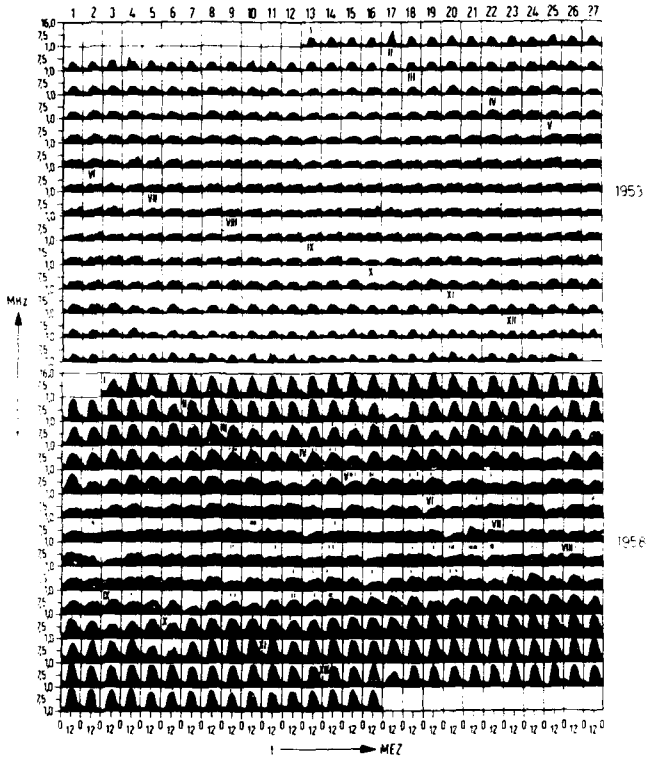



Fig. 3. "Lindau Mountain Diagrams" showing the critical frequencies foF2 at [illegible] for the sunspot minimum year 1953 and the sunspot maximum year 1958. Each box represents [illegible] MHz; the data are arranged in lines of 27 days in order to display recurrences [illegible] 27-day rotation period. Copyright: Max-Planck-Institut für Aeronomie, Fallenbostel, [illegible] Republic.

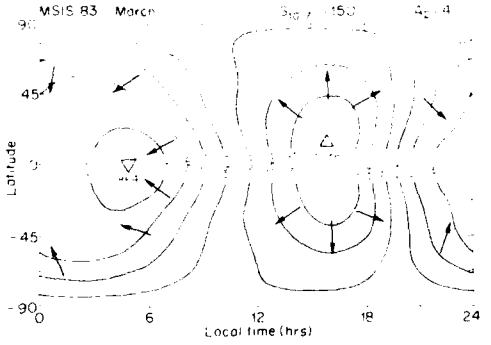



Fig. 4. Map of exospheric temperature at March equinox, according to the MSIS83 thermospheric model [4], for quiet magnetic conditions, high solar activity. Arrows show approximate directions of the thermospheric winds.

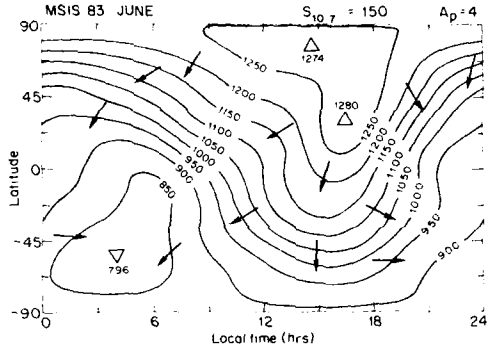



Fig. 5. As Fig. 4, for June solstice.

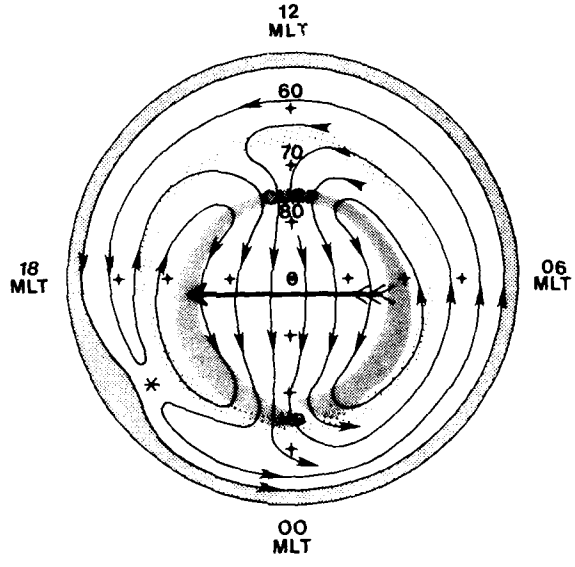



Fig. 6. The auroral oval and associated phenomena. View of the north polar region, centred on the magnetic pole, in a non-rotating frame of reference. Magnetic midnight (00 MLT) is at the bottom; magnetic noon (12 MLT) at the top; the outer boundary is at 50 deg magnetic latitude; magnetic latitudes 80, 70, 60 deg are marked by crosses. The outer ring represents the plasmasphere (mid-latitude ionosphere) which co-rotates with the Earth. The shaded ring represents the approximate location of the hard "drizzle" precipitation. The inner shading is the auroral oval, showing the approximate locations of the "cusp" in the noon sector and the Harang discontinuity (HD) near midnight. The fine lines represent typical flow lines of the plasma convection pattern (some curves are left incomplete to reduce congestion); note the stagnation point at 20 MLT. The diagram corresponds roughly to moderate magnetic activity (K = 3) with a southward *Bz* component of the interplanetary magnetic field.
Copyright: S.E.R.C. Rutherford Appleton Laboratory, U.K.

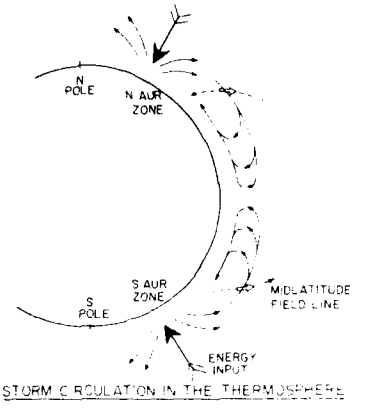



Fig. 7. Sketch of the "thermospheric storm circulation" and its effects in the F region.

## DISCUSSION

**K.Rawer**

Original ionospheric predictions were for a long time in advance using observed diurnal, seasonal, and solar cycle variations to give a monthly average prediction. The uncertainty is a few MHz in $f_0F_2$ when applying such a prediction to an individual day. (That system might be improved at low latitudes by introducing the main effects of MHz of the magnetic equator). Now, considering a theoretically based prediction method, we may compare with what is now done in meteorology. Meteorologists start with an empirically obtained hemisphere wide pattern describing the actual state. Then, they use theory to calculate the future development of the system. This method has two shortcomings:

(1) Sometimes the decisive cell is not covered by the input data; quite small regions might become decisive for the future development, unfortunately.

(2) What one might call "noise" is not only present in the input data, but occurs also during the development of the system. Noise can be any effect which, although of physical origin, cannot be introduced into the schedule. This has for effect that such meteorological predictions can – at least at present – only be extended over a period of about one week at the maximum. Similar ionospheric prediction efforts may encounter comparable problems.

**Author's Reply**

(Regarding 2) There is of course a great deal of "noise" in ionospheric measurements, as in all other branches of geophysics. Ionospheric physicists have to use their skill and intuition to extract "physics" from the "noise".

This leads on to the question of maintaining an adequate network of ionosondes to monitor the ionosphere. A recent map shows about 100 active sounders, but the number is decreasing; the geophysically important station in New Zealand has just closed. Not every existing station will be needed forever, but there should be a way of determining what network is really needed, and finding a way of supporting it. The EPP should consider this question.

**P. Vila**

(1) The low-latitude "picture" of storm effects on $f_0F_2$ needs separating *dip Equator* changes, giving storm increases from tropical latitude ones where most storm conditions are a large decrease. (Unless dip equator converges).

(2) About the $O/N_2$ "non-change" suggested following auroral disturbances, it seems from suggested St. Santin results that the Prölss model is partially correct, to give a change in $O/N_2$ at the $f_0F_2$ level.

**Author's Reply**

No comment on (1).

(2) I agree that the $O/N_2$ ratio does seem to decrease at the midlatitude $F2$ peak during storms; that is quite well documented. The UCL model does not give this decrease. A possible explanation is an input of energy from the ring current.

**B.Showen**

Prof. Rishbeth, you wrote a paper a few years ago with Ganguli and Walker, modeling $h_mF_2$ and $N_mF_2$ variations and comparing them with Arecibo and St. Santin data. Is your model still a good one, and has more recent work been done with it?

**Author's Reply**

The model, although quite simple, is still valid. It used the reduced height method, although that is not implicitly stated in the paper.

# A FULLY ANALYTIC, LOW AND MIDDLE LATITUDE IONOSPHERIC MODEL

D.N. Anderson and J.M. Forbes*
Air Force Geophysics Laboratory
Hanscom AFB, MA 01731-5000

## SUMMARY

The well-known Chiu ionospheric model is a global empirical model which calculates electron density profiles as a function of latitude, local time, season and solar cycle. It is extensively used by thermospheric dynamicists because it's analytic formulation can quickly provide global electron density values. The model, however, has serious shortcomings. Recently, a fully analytic ionospheric model based on the Chiu formulation, has been developed which provides much more realistic, low and mid-latitude electron density distributions. This is accomplished by retaining the latitude and local time dependence of parameters such as hmF2, nmF2, and the topside and bottomside scale heights. The modifications are based on a recently-developed Semi-empirical, Low-latitude Ionospheric Model (SLIM) and involves decomposing hmF2, nmF2 and plasma density scale heights into harmonic terms to describe the local time variations and then applying Hermite polynomials to obtain latitude variations. The analytic model generates electron density profiles under solar cycle minimum, moderate and maximum conditions for winter, summer and equinox periods. Differences between this model and the Chiu model are discussed and comparisons with observed electron density distributions are presented.

## INTRODUCTION

Recently, a Semi-empirical, Low-latitude Ionospheric Model (SLIM) was developed (Anderson et al., 1987a) which provided Faraday electron content values from 150 km to 1000 km every 2° latitude (-24° to 24°) and every half hour over the 24-hour day. The reason for producing this data was to provide accurate electron density profiles which were more realistic than an existing model, which is the Chiu (Chiu, 1975) or Bent (Llewellyn and Bent, 1973) model. To separate these profiles in a computationally efficient manner,

More briefly, the SLIM approach was to find functional fits to empirically theoretically-calculated electron density profiles. These profiles are composed of three basic sets of numbers which represent the profile parameters corresponding to the three peaks, E-valley, [illegible] and Topside for three levels of solar activity, solar minimum, [illegible] and solar maximum, $I_{pp} = 250$ sfu for each peak. [illegible] six parameters, $nmF2$, $hmF2$, $B_{bp}$, $B_{tp}$, [illegible] and the [illegible]

[illegible]

where $nE2$ is the peak plasma density and $\frac{h-hmF2}{B}$.

A separate set of six coefficients exist for each 2° interval in latitude and 1 hr interval in local time. While electron density profiles can be calculated from the various sets of coefficients every 2° in latitude, it was found that [illegible] and latitude, an approach adopted in the Chiu empirical model. [illegible] a theoretical model has been extensively used to specify low and mid-latitude ionospheric fields, it was decided [illegible] the coefficients of fits to the calculations. The [illegible] model will [illegible] to the user, [illegible] a set of curves to modify and [illegible] local and solar [illegible].

This paper describes Bent's [illegible] that are used to generate the [illegible] which will [illegible] the [illegible] and then [illegible] Ionosphere Model Institute [illegible] model [illegible] illustrate the [illegible] and [illegible] to show comparisons between [illegible] and [illegible] with [illegible] calculating the [illegible] time [illegible] world [illegible] with [illegible].

## THE MODEL

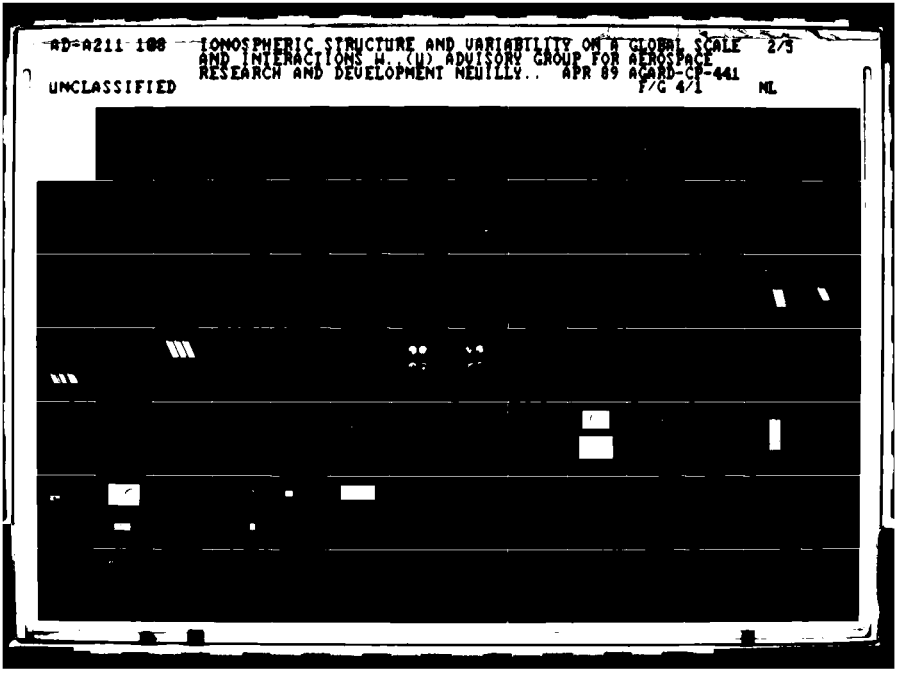
The empirical Chiu ionospheric model is based on middle and low latitude observations of the E, F1, and F2 regions of the ionosphere over the [illegible]

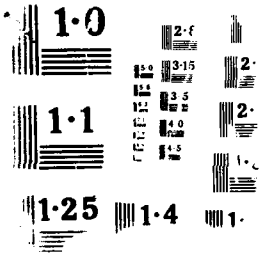
---

* Major, USAFR. Permanent Affiliation: Center for Space Physics, Boston University, Boston, MA 02215

AD-A211 108 IONOSPHERIC STRUCTURE AND VARIABILITY ON A GLOBAL SCALE 2/3
AND INTERACTIONS W..(U) ADVISORY GROUP FOR AEROSPACE
RESEARCH AND DEVELOPMENT NEUILLY.. APR 89 AGARD-CP-441

UNCLASSIFIED F/G 4/1 NL




1·0
1·1
1·25
1·4

variations as well as the solar cycle variations. The model assumes that the F2-region plasma density can be expressed by a modified Chapman expression given by

$$N_e(H) = nmF2\exp[1-Z-\exp(-Z)] \qquad (2)$$

where nmF2 is the peak electron density, and $Z = \frac{h-hmF2}{H}$; hmF2 is the peak altitude of the F2 layer where $N_e = nmF2$ and H is an appropriate plasma density scale height. In the Chiu formulation, two expressions for Z are specified, one for the bottomside of the F layer and one for the topside given respectively by

$$Z_{LO} = \frac{h - hmF2}{H_{LO}}$$

$$Z_{UP} = \frac{h - hmF2}{H_{UP}}$$

In all, four coefficients are needed to completely specify the F2-layer in the Chiu model, nmF2, hmF2, $H_{UP}$ and $H_{LO}$. Similar formulations are utilized for the F1- and F-layers, and the total plasma density as a function of altitude is obtained by summing these three contributions. Analytic functions are used throughout to express the coefficients as a function of geographic and geomagnetic coordinates, sunspot number, local time, and day of year.

The Chiu model is widely known to be deficient in its specification of hmF2 at middle and low latitudes, and its representation of the equatorial anomaly. In our approach we have first applied a correction to the Chiu hmF2 values at middle latitudes based on ionosonde data. Then, we have added analytic corrections for foF2 and hmF2 and modified $H_{UP}$ and $H_{LO}$ in the Chiu model between ±24° dip latitude such that the ionization profiles best fit those of the SLIM model. This procedure and the equations involved are detailed below.

## ANALYTIC FORMULATION

The middle latitude correction to the hmF2 specification of the Chiu model is based on examination of equinox ionosonde data, as detailed in an AFGL report (Forbes et al., 1988) which can be provided to the reader upon request. The original expression for hmF2 in Chiu (1975) is as follows:

$$hmF2 = 240 + 75\sigma + 83\sigma\zeta\cos\lambda_m + 30\cos(\phi - 4.5|\lambda| - \pi) + 10\cos\lambda_m\cos\tfrac{\pi}{3}(t-4.5)$$

where $\sigma$ = R/100 where R is the monthly smooth Zurich sunspot number,
$\lambda_m$ = geomagnetic latitude (rad),
$\zeta = \sin\delta\sin\lambda_m$ where $\delta$ = solar angle (rad),
$\phi$ = local time angle (rad) measured from local midnight.

Note that the local time variation consists only of a diurnal (24-hour harmonic) which is independent of season and solar cycle. Our correction concerns the second (solar cycle) and fourth (local time) terms in the above equation. The former is replaced by

$$100\sqrt{\sigma} - 10$$

The local time term is replaced by

$$(56-12\sqrt{\sigma})\cos(\phi + \lambda_\alpha) + (9 + 12\sqrt{\sigma})\cos(2\phi - \tfrac{\pi}{4}) + 12\cos(3\phi - \tfrac{3\pi}{4})$$

where $\lambda_\alpha = -2\pi + \frac{3\pi}{4}e^{-(\frac{\lambda_m}{20})^2}$.

In developing the low-latitude portion of FAIM, we derive new analytic formulations for the parameters foF2, hmF2, $H_{UP}$, and $H_{LO}$ in the Chiu model based on least-squares fits to the SLIM profiles. The local time variation at a given dip latitude ($\lambda_m$) is expressed as a series of sines and cosines with amplitudes $a_n$ and $b_n$ with the series going from n = 0 to n = 6:

$$f(\lambda_m,\phi) = \sum_{n=0}^{6}(a_n(\lambda_m)\sin(n\phi) + b_n(\lambda_m)\cos(n\phi)) \qquad (3)$$

where $\phi$ is the local time from midnight in radians.

In reality, for foF2 and hmF2 values, we used the expression above to fit the difference between the Chiu and SLIM values so that in the new analytic model, we simply calculate Chiu values and then add to these the analytically-formulated difference values to provide the final foF2 and hmF2 quantities specified by FAIM.

The procedure for obtaining the topside and bottomside F-region scale heights, $H_{UP}$ and $H_{LO}$, is handled somewhat differently. Here, a separate topside and bottomside scale height was found which would reproduce each SLIM density profile to within 0.5 MHz in plasma frequency assuming the expression for $N_e(h)$ given by Eq. (2). These scale heights are different from those specified in the paper by Anderson et al. (1987a) because the values for $C_{UP}$ and $C_{LO}$ are now assumed to be unity. Once values of $H_{UP}$ and $H_{LO}$ were obtained by this method, the analytic function given by Eq. (3) was used to derive $a_n$ and $b_n$ values.

The next task was to find appropriate functions and coefficients which would reproduce the separate sets of $a_n$'s and $b_n$'s as a function of dip latitude ($\lambda_m$). The modified harmonic oscillator function $H_k(\tau)e^{-\tau^2}$ where $H_k$ is the Hermite Polynomial of order k and $\tau = \frac{\lambda}{14}$ was found to be ideally suited for this purpose:

$$a_n(\lambda_m), b_n(\lambda_m) = \sum_{k=0}^{4} C_k H_k(\tau) e^{-\tau^2} \qquad (4)$$

The Forbes et al. (1988) report contains a tabulation of the $C_k$'s as well as the FORTRAN code for FAIM. The code can also be supplied to requestors over SPAN, BITNET, or on floppy disk.

## RESULTS

As outlined in the paper by Anderson et al. (1987a) nine separate SLIM cases were generated which covered three levels of solar activity, low, moderate, and high for each of three seasons, equinox, June solstice, and December solstice. For these nine cases, sets of coefficients to the analytic functions described above were determined which would reproduce the local time and latitude variations in the four parameters foF2, hmF2, $H_{UP}$, and $H_{LO}$.

To demonstrate how well FAIM reproduces the SLIM values, we have plotted in Figure 1 the latitude variation in hmF2 and foF2 at 1400 LT for the solar cycle maximum, equinox case. In this figure, we also compare these models with the empirical Chiu values for hmF2 and foF2. Looking first at the hmF2 comparisons, FAIM smooths out the SLIM values between 12°N and 12°S dip latitude and gives excellent agreement poleward of ±12°. Actually, the large variation in SLIM hmF2 between ±4° and ±10° dip latitude is somewhat an artifact of the SLIM least-squares approach and does not represent reality so that the smoother variation in FAIM hmF2 values is more realistic. The comparison of both SLIM and FAIM with the Chiu model illustrates the magnitude of the improvement at low latitudes. As reported by Rajaram and Rastogi (1977) daytime hmF2 values near the magnetic equator at Kodaikanal, India lie in the range 450 to 500 km, about 100 km greater than the 350 km specified in the Chiu model.

The comparison in foF2 between FAIM and SLIM is excellent. The analytic model slightly overestimates SLIM values near the crests of the equatorial anomaly at ±10° to ±20° dip latitude and underestimates slightly near the magnetic equator, but these differences are well within the day-to-day fluctuations in equatorial foF2 values. A comparison with the Chiu model shows the crest-to-trough density ratio is greater for the analytic model which is in good agreement with observed values in the Indian sector during March 1959, a solar cycle maximum equinoctial period (Rastogi, 1966).

At 2000 LT, the comparisons between SLIM, FAIM, and Chiu are even more impressive as shown in Figure 2. The post-sunset enhancement in upward ExB drift is a characteristic feature during equinox, solar maximum years (Woodman, 1970) and is responsible for the very high hmF2 values near the magnetic equator (Anderson, 1973) and is reproduced in both SLIM and FAIM. Such high hmF2 values are also in agreement with measured values at Kodaikanal (Rajaram and Rastogi, 1977) and at Jicamarca (Anderson et al., 1987b). The analytic model reproduces the latitude variation in SLIM values very well and both these models give substantially greater hmF2 values than the Chiu model.

The enhancement in the upward ExB drift velocity at the equator mentioned above is also responsible for maintaining the crests of the anomaly out to ±18° dip latitude and increasing the peak-to-trough ratio after sunset. This latitude variation in foF2 at 2000 LT is completely absent in the Chiu model which essentially displays a constant density value with latitude. The more realistic FAIM and SLIM models with peak densities of $5 \times 10^6$ el/cm$^3$ (foF2 ~ 18 MHz) at the crests of the anomaly, provide much greater ion drag forces at these latitudes which translates to a greater latiude variation in the zonal neutral wind velocities. In addition, this latitude structure in peak densities significantly affects flux-tube-integrated Pedersen conductivities after sunset, which play an important role in the electric coupling between the low latitude ionosphere and neutral atmosphere (Anderson et al., 1987b).

The unrealistically low F layer height provided by the Chiu model has significant

consequences if the ionospheric model is incorporated into a dynamic neutral atmospheric model. After sunset, near the equator, the two main horizontal forces on the neutrals are the pressure gradient and ion drag forces, and if the ionospheric densities are unrealistically large near 250 km (as they would be if hmF2 is underestimated), then calculated zonal wind speeds will be significantly underestimated (Anderson and Roble, 1974).

Comparisons of profile shapes at the magnetic equator and at 16°S dip latitude are illustrated in Figures 3 and 4 for the 2000 LT period. At the magnetic equator, reasonable agreement between FAIM and SLIM is achieved with the topside portion being in better agreement than the bottomside region. Actually, this represents one of the poorest fits of all the comparisons. The degree to which both FAIM and SLIM profiles differ from the Chiu model is graphically illustrated and displays the unrealistically large electron density values in the 350 km region provided in the Chiu model. At 16°S dip latitude, the comparison between FAIM and SLIM is excellent and is more indicative of the general agreement which can be achieved when analytic functions for $H_{UP}$ and $H_{LO}$ are used to calculate profile shapes.

## 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., 21, 421-442, 1973.

Anderson, D.N. and R.G. Roble, The effect of vertical ExB ionospheric drifts on F-region neutral winds in the low-latitude thermosphere, J. Geophys. Res., 79, 5231-5236, 1974.

Anderson, D.N., M. Mendillo, and B. Herniter, A semi-empirical, low-latitude ionospheric model, Radio Sci., 22, 292-306, 1987a.

Anderson, D.N., R.A. Heelis, and J.P. McClure, Calculated nighttime eastward plasma drift velocities at low latitudes and their solar cycle dependence, Annal. Geophys., A5, 435-442, 1987b.

Chiu, Y.T., An improved phenomenological model of ionospheric density, J. Atmos. Terr. Phys., 37, 1563-1570, 1975.

Forbes, J.M., D.N. Anderson, P. Batista, and M. Codrescu, An analytic/empirical model of the middle and low latitude ionosphere, AFGL Technical Report, (in press), 1988.

Llewellyn, S.K. and R.B. Bent, Documentation and description of the Bent ionospheric model, Rept. AFCRL-TR-73-0657, AD 772733, Air Force Cambridge Research Laboratories, Bedford, MA, 1973.

Rajaram, G. and R.G. Rastogi, Equatorial electron densities - seasonal and solar cycle changes, J. Atmos. Terr. Phys., 39, 1175-1182, 1977.

Rastogi, R.G., The equatorial anomaly in the F2 region of the ionosphere, Inst. Telecomm. Engrs., 12, 245-256, 1966.

Woodman, R.F., Vertical drift velocities and east-west electric fields at the magnetic equator, J. Geophys. Res., 75, 6249-6259, 1970.

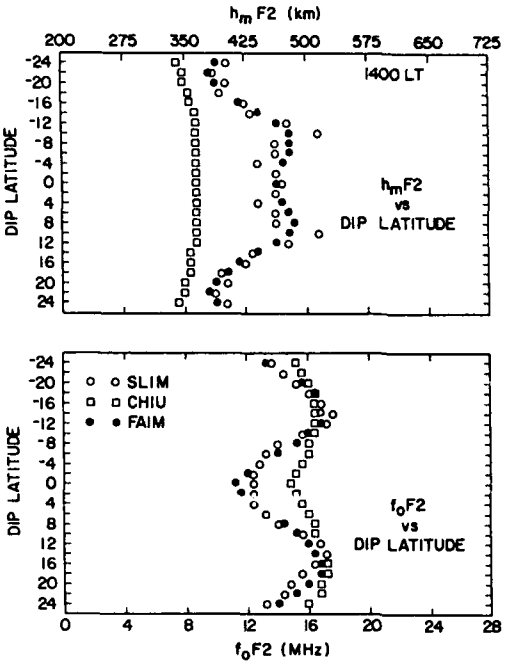
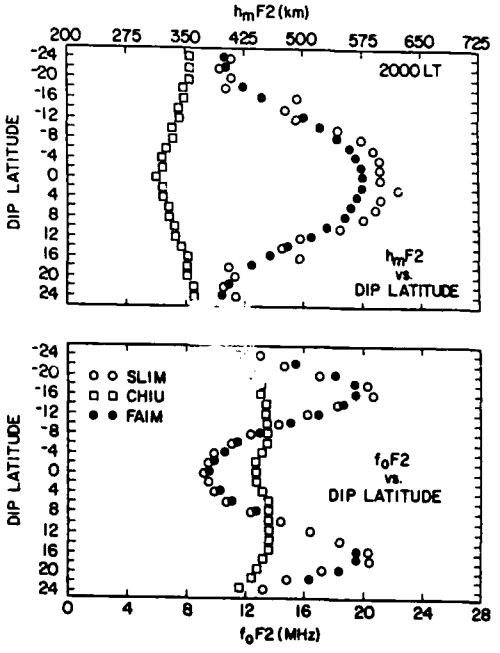



FIGURE 1. Comparison of hmF2 and foF2 values as a function of dip latitude at 1400 LT given by *SLIM*, *CHIU* and *FAIM* for solar cycle maximum, equinoctial conditions

FIGURE 2. Same as in Figure 1 except at 2000 LT



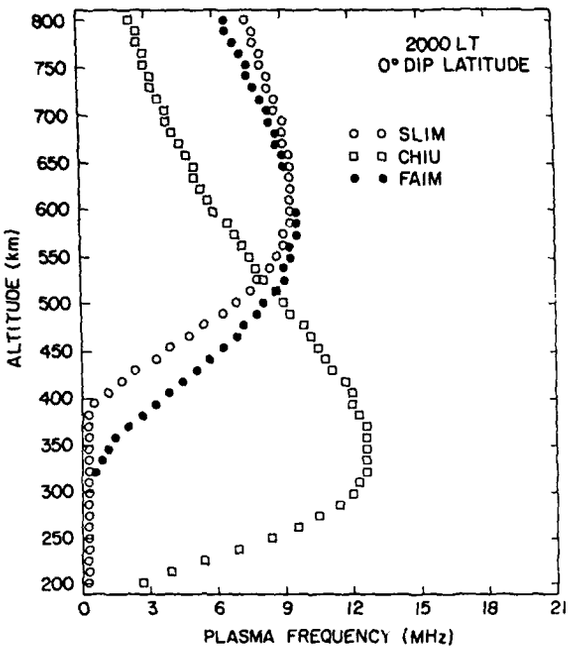



FIGURE 3. Comparison between SLIM, CHIU and FAIM electron density profiles at the magnetic equator at 2000 LT

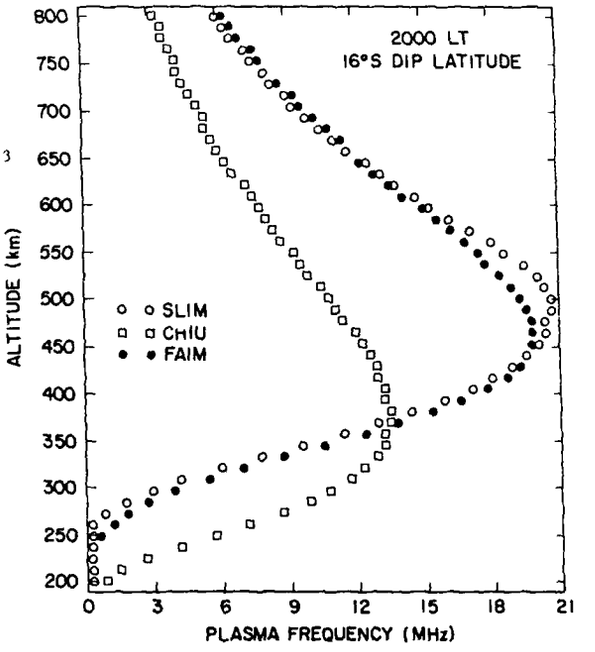



FIGURE 4. Same as in Figure 3 except at 16°S dip latitude

## DISCUSSION

**K.Rawer**

The shortcoming of IRI with respect to layer thickness in certain geographical regions is due to the lack of input data at the time when the model was built. We had only two low-latitude stations; only the northern middle-latitude region was well covered (Germany). Following a proposal by Madam Gulyaeva, we do now try to get more data from true height analysis. The parameter proposed is the height where half the peak density is found. We hope to be able to convince a few station networks to go ahead in determining this new parameter. We need that because empirical models should be built upon measured data.

**Author's Reply**

I would note that often in the ocean and high-latitude regions you cannot even get empirical ground-based data, and therefore, maybe you need to resort to theoretical models to get the realism you need.

**P.Vila**

On the latitude profiles from the SLIM model, there are two small troughs on either side of the dip equator. Are these real or are they numerical errors?

**Author's Reply**

They are numerical errors.

# GLOBAL MORPHOLOGY OF PLASMA BUBBLES IN THE LOW LATITUDE IONOSPHERE

by

K. Schlegel[1], K. Oyama[2], T. Takahashi[3]

[1] Max-Planck-Institut für Aeronomie, Katlenburg-Lindau, FRG

[2] Institute of Space and Astronautical Science, Tokyo, Japan

[3] Tohoku University, Sendai, Japan

## ABSTRACT

The electron density and electron temperature structure in and around plasma bubbles in the low latitude ionosphere has been studied with instruments on the Japanese satellite HINOTORI. Histograms are presented about the number of occurrence of certain types of bubbles in a latitude range of ±30° around the equator. Particularly the temperature within the plasma bubbles which can be higher or lower than the temperature of the ambient plasma will be discussed and explanations will be proposed. A particular class of bubbles seem to occur over the South Atlantic magnetic anomaly and over the Hawaian anomaly.

## 1. ITRODUCTION

The term "plasma bubble" was coined by Woodman and LaHoz (1976). They explained large plumes raising in the night time equatorial ionosphere, which they observed with the Jicamarca radar fecility, as an analogon of lighter fluid bubbles raising in a heavier fluid. Hanson and Sanatani (1973) who observed the same phenomenon earlier with a space-borne retarding potential analyzer, called them "biteouts", since they found the ion density reduced by up to three orders of magnitude inside these plasma voids. According to McClure et al. (1977) the main features of these bubbles can be summarized as follows: They are probably cylindrical in shape with the cylinder axis aligned with the magnetic field. Their scale size is of the order of several tens to a few hundred km, and they generally move upward with a velocity of a few 100 m/s. Often a slower (< 100 m/s) westward motion with respect to the ambient plasma is observed. Ion density reductions of a factor of 1000 as quoted above seem to be extreme cases, mostly the reduction is of the order of 10-100.

Plasma bubbles should be clearly distinguished from plasma blobs, which are localized density enhancements (Watanabe and Oya 1986).

## 2. SATELLITE INSTRUMENTATION

The 7th scientific Japanese satellite 'Hinotori' was launched on the 21st of February 1981 to observe the solar surface during high solar activity. Hinotori was a spinning satellite which was controlled in a way that the spin axis always pointed towards the sun. Two plasma probes were installed on the satellite, in order to monitor the ionosphere in addition to the solar observations. One was an *electron temperature probe (Oyama and Hirao, 1982)* and the other a *swept frequency* impedance probe for electron density measurements (Oya and Morioka, 1975). The sensor electrodes of these two experiments were symmetrically mounted at the edge of 2 of the 4 solar cell paddles of the satellite. The sensor of the electron temperature probe was a circular disk of 120 mm in diameter and 1.6 mm thick, while the sensor of the electron density probe was cylindrical in shape with a diameter and length of 10 mm and 400 mm, respectively. Signals from both these two plasma probes were more or less modulated by the satellite spin. The degree of the data modulation depended on the location of the satellite. The strongest signal modulation occurred around midnight and noon where the spin axis of the satellite was perpendicular to the direction of the satellite motion. Data are gathered when the sensor plane of the electron temperature probe was in the ram direction with respect to the satellite motion using the fact that when the electron density sensor was in the satellite wake, the electron temperature probe was out of the wake. The sensor impedance of the electron density probe showed a clear minimum when the probe was in the wake of the satellite. Therefore the data resolution was determined by the satellite spin rate, and the horizontal scale resolution along the satellite pass was approximately 128 km. The electron density probe and the temperature probe detected 724 plasma bubbles from 4000 orbits during the period from June 1981 to February 1982. The observations were limited to geographic latitudes of ± 30 degrees. The data obtained during one satellite pass around the earth were stored in an onboard data recorder and recovered when the satellite was overhead Kagoshima Space Center.

## 3. GENERAL BEHAVIOR OF PLASMA BUBBLES

*Figure 1 shows a typical example of two plasma bubbles separated by about 250 km.* The density depletion is of the order of 100 in both cases. These bubbles are very typical for the evening topside ionosphere. The bubbles can have internal structure like the two shown in the figure, also several bubbles can appear in a series.

The latitudinal distribution of bubbles is shown in figure 2. It is obvious that they are symmetrically distributed around the magnetic dip equator, and that their occurence probability maximizes around equinox. It should be noted that the occurrence probability is the ratio of observed bubbles in a certain time interval to the actual available observation interval.

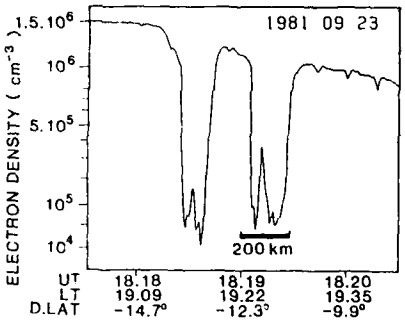



**Figure 1.** Typical example of plasma bubble profiles observed by the HINOTORI satellite.

Figure 3 shows the occurrence probability versus local time indicating that the bubbles are *mainly a night-time* phenomenon. The structure of these distributions is statistically significant, at least for the equinox periods. The two peaks in this distribution, one around 22 LT, the other around 2 LT probably correspond to different types of bubbles. Oya et al. (1986) distinguish between the so called "multiple plasma bubbles" (MPB) and the "quasi periodic plasma bubbles" (QPB). The MPB, a series of bubbles observed along the satellite trajectory occur predominantly in the pre midnight period. They are most likely associated with a Rayleigh-Taylor type instability mechanism (Hærendel 1974). Figure 3 shows a sharp increase of the bubble occurrence probability immediately after 18 LT. This time corresponds at equinox to a solar zenith angle of 105° which means sunset at about 300 km altitude. The steepening of the electron density profile at the bottom of the F-region after this time can be the cause for the above mentioned instability. The lifetime of the bubbles, slowly drifting upwards, can be several hours.

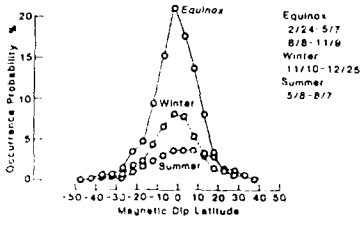



**Figure 2.** Occurrence probability of equatorial plasma bubbles as a function of magnetic dip latitude for three different seasons.

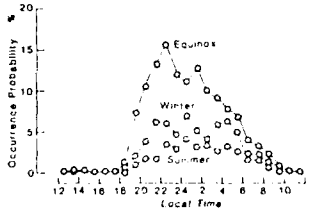



**Figure 3.** Same as figure 2 but as a function of local time.

A different production process is assumed for the QPB which occur mainly after midnight. Their name stems from the fact that these bubbles are associated with wavelike structures in the background electron density. They can often be traced over several satellite trajectories. Probably they are created in responce to dynamical disturbances of the background ionosphere. These perturbations can either be gravity waves or magnetic disturbances. The fact that a correlation between the geomagnetic activity and the occurence of QPB was found (Watanabe and Oya 1986) supports this interpretation. With increasing Kp the occurrence of QPBs increases in the morning hours.

It should be noted that these two types of bubbles are apparently related to two types of equatorial spread F described earlier by Röttger (1975). The so called "early spread F" corresponds to MPB, the "late spread F" to QPB.

The seasonal variation of the occurrence probability is still not fully understood. It probably reflects the seasonal variation of the control parameter of the corresponding plasma instabilities. Substorms (Takahashi et al. 1987), the global wind distribution (Chiu and Strauss 1979) and more localized weather phenomena may also effect the appearance of bubbles. It is interesting to note that during northern summer and around equinox the bubbles appear mainly above Africa, while during the northern winter the maximum of occurrence is located over South America (Watanabe and Oya 1986). This suggests a relation to thunderstorm activity (Röttger 1977). A further discussion of the seasonal variation is beyond the scope of this paper.

## 4. ELECTRON TEMPERATURE STRUCTURE OF THE BUBBLES

The electron temperature inside the bubbles is difficult to measure. Since the electron density can change by more than two orders of magnitude within a few seconds when a bubble is traversed, the dc amplifiers have to step through several gain levels which requires a fast switching and a high data rate. Therefore electron temperature results in and around bubbles were scarce before the Hinotori measurements (see Oyama et al. 1988 for references).

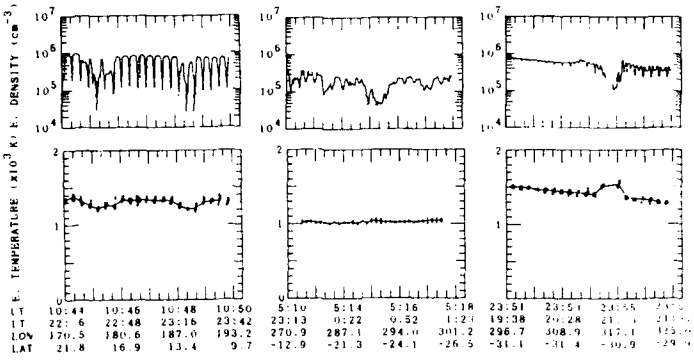



**Figure 4.** Three different types of electron temperature behaviour (lower panels) in and around plasma bubles (upper panel): The temperature inside is lower (left), equal (middle), and higher then outside (right). The density fluctuations seen in the upper left panel is due to the satellite wake effect.

Three different types of electron temperature behaviour were found in the data, they are displayed in figure 4. The left hand panels show a case where the bubble's inner temperature (denoted by $T_i$ in the following) is smaller than the outside temperature ($T_o$). In the middle panels both temperatures are equal, and in the right hand panels $T_o > T_i$. Before giving possible explanations for these three cases, the temperature difference $\Delta T = T_i - T_o$ should be examined in more detail.

Figure 5 shows the local time dependence of this temperature difference. It is obvious that in most of the cases the inner temperature of the bubbles is smaller than the temperature outside. A fair number of bubbles, however, exist where the inner temperature can be substantially higher than the outside temperature, mostly in the morning hours but some occur also in the evening. Around 0300 LT, however, the inner and the outer temperature are almost equal.

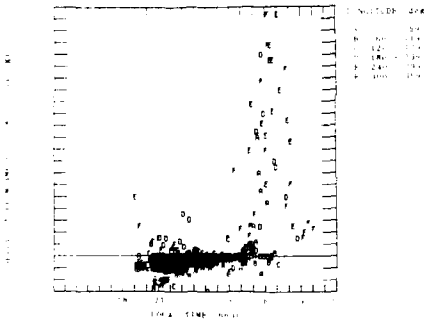


**Figure 5.** $\Delta T$ = (inner temperature − outer temperature) versus local time. The letters indicate the longitude range in which the bubbles were observed.

A probable explanation of the high positive $\Delta T$ values in the morning is photoelectron heating. It can be shown that in almost all these cases the bubbles are still in the dark part of the ionosphere, while the height region above is already sunlit. Thus photoelectrons are produced above the bubbles, spiraling down the magnetic field line and enter the bubbles. Due to the conservation of energy these photoelectrons heat the density depleted bubbles stronger than the ambient electron gas. In addition to the photoelectrons produced in the same hemisphere, photoelectrons from conjugate points may play a role when the opposite hemisphere is earlier sunlit than the observation region.

A different explanation exist for the about 20 cases with significant positive $\Delta T$ between 20 and about 2 LT. A close examination shows that these bubbles are mainly concentrated above two geographical regions. In Figure 6 the location of these bubbles is plotted in a world map, where only bubbles with $\Delta T \geq 50$ K (50 K being roughly the accuracy of the measurements) are taken into account. The majority of points lay in two wellknown magnetic anomaly regions, the South Atlantic Anomaly and the Hawaian Anomaly. Due to the special geometry of the magnetic field lines, energetic electrons from the magnetosphere can precipitate in these regions (e.g. Voss and Smith 1980, Varga et al. 1985), and may heat up the thermal electron gas inside the bubbles. It should be noted that anomalously high electron temperatures have been observed in these regions also in the absence of bubbles, probably caused by the same mechanism (Oyama and Schlegel 1984).

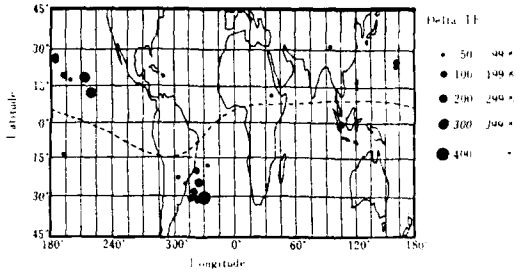



**Figure 6.** Location of plasma bubbles with $\Delta T > 50$ K occurring in the pre-midnight period. The radius of the circles indicates the magnitude of $\Delta T$. A dashed line shows the magnetic equator.

It can be shown by comparing the diffusion time perpendicular to the magnetic field lines and the raise time of the bubbles that thermal electrons from outside are not able to enter the bubble during its path upwards into the topside ionosphere (Oyama et al. 1988). Therefore a heating of the bubble can only be caused by particles moving along the geomagnetic field lines. This is the case for the two categories discussed above, in one case photoelectrons in the other case energy electrons.

The fact that the majority of bubbles observed after sunset (around 21 LT in figure 5) exhibit a positive $\Delta T$ is easy to explain. At these times the electron temperature profile in the ionosphere still has a positive slope. Thus bubbles generated at altitudes below 600 km should have lower temperatures inside when they finally reach the equatorial altitudes. Later in the night the electron temperature becomes more and more constant with height and thus $T_e$ in the bubble appears hotter in agreement with the results shown in figure 5.

## CONCLUSION

While there are still several open questions in the production mechanism and the seasonal distribution of plasma bubbles at low latitudes their thermal behaviour is relatively well understood. It should finally be noted that the occurrence of bubbles in the night time low latitude ionosphere imposes some implications on the use of this part of the ionosphere as a radio wave transmission medium. Three different phenomena have been observed which are related to the bubbles in the radio propagation language called "equatorial spread F":

i. A deviation from the usual transequatorial great circle propagation of HF signals due to scattering of the waves by the density structures of the bubbles. Azimuth deviations of more than 20° have been observed on a transmission path between stations in Germany (Lindau) and South Africa (Tsumeb) (Röttger 1972).

ii. Strong fading of the received amplitudes with a broad spectral distribution, so called flutter fading (Röttger 1972). It is similarly caused by the varying density and the motion of the bubbles.

iii. Occasionally transequatorial propagation of VHF waves (50 - 150 MHz) over more than 6000 km. This is caused by forward scattering of the waves or more accurately by ducting in the field aligned density depletions, called bubbles (Röttger 1979).

One can therefore state that the in situ investigations of these plasma bubbles by satellite techniques complement the earlier results obtained by radio propagation and scintillation (Basu and Basu 1981) experiments.

### Acknowledgements

The authors would like to express their sincere thanks to the satellite launching staff of the Institute of Space and Astronautical Science, Tokyo.

## REFERENCES

Basu S. and Basu S., Equatorial scintillation - A review, J. Atmos. Terr. Phys. **43**, 473-489, 1981.

Chiu Y.T. and Strauss J.M., Rayleigh-Taylor and wind driven instabilities of the night time equatorial ionosphere, J. Geophys. Res. **84**, 3283-3290, 1979.

Haerendel G., Theory of equatorial spead F, Report, Max-Planck Institut für Phys. and Astrophys., Garching, F.R.G., 1974.

Hanson W.B. and Sanatani S., Large $N_i$ gradients below the equatorial F peak, J. Geophys. Res., **78**, 1169-1193, 1973.

McClure J.P., Hanson W.B., and Hoffman J.H., Plasma bubbles and irregularities in the equatorial ionosphere, J. Geophys. Res., **82**, 2650–2656, 1977.

Oya H. and Morioka A., Instruments and observations of gyro-plasma probe installed by Taiyo for measurement of ionospheric plasma parameters and low energetic particle effects, J. Geomag. Geoelectr. **27**, 331-361, 1975.

Oya H., Takahashi T. and Watanabe S., Observation of low latitude ionosphere by the impedance probe on board the HINOTORI satellite, J. Geomag. Geolectr. **38**, 111-123, 1986.

Oyama K.-I., and Hirao K., Electron temperatures probe on board Japan's 7th Scientific satellite Hinotori, Proc. 13th ISAS Symp, 1982.

Oyama K.-I. and Schlegel K., Anomalous electron temperatures above the South American magnetic field anomaly, Planet. Space Sci., **32**, 1513–1522, 1984.

Oyama K.-I., Schlegel K. and Watanabe S., Temperature structure of plasma bubbles in the low latitude ionosphere around 600 km altitude, Planet. Space Si. 1988, in press.

Röttger J., Some effects of atmospheric gravity waves observed on a transequatorial radio path, AGARD Conf. Proc. **115**, 33-1 - 33-15, 1972.

Röttger J., Influence of spread-F on HF radio systems, AGARD Conf. Proc. **173**, 26-1 - 26-19, 1975.

Rottger J., Travelling disturbances in the equatorial ionosphere and their association with penetrative cumulus convection, J. Atmos. Terr. Phys. **39**, 987-998, 1977.

Röttger J., The phenomenology of transequatorial radio propagation under spread-F conditions, AGARD Conf. Proc. **263**, 24-1 - 24-9, 1979.

Takahashi T., Oya H. and Watanabe S., Ionospheric disturbances induced by substorm associated electric fields in the low latitude F-region, J. Geomag. Geoelectr. **39**, 187-209, 1987.

Varga L., Venkatesan D. and Meng C.-I., Low altitude observations of the energetic electrons in the outer radiation belt during isolated substorms, Planet. Space Sci., **33**, 1259–1266, 1985.

Voss H.D. and Smith L.G., Global zones of energetic particle precipitation, J. Atmos. Terr. Phys. **42**, 227-239, 1980.

Watanabe S., and Oya H., Occurence characteristics of low latitude ionosphere irregularities observed by impedance probe on board the Hinotori satellite.
J. Geomag. Geoelectr., **38**, 125-149, 1986.

Woodman R.F., and LaHoz C., Radar observations of F region equatorial irregularities, J. Geophys. Res., **81**, 5447-5466, 1976.

## DISCUSSION

**R.F.Benson**

When using satellite observations to investigate plasma bubbles one must use care to make sure that large depletions attributed to bubbles are not in reality due to the penetration of the bottomside of the $F$ region by the satellite. Studies based on DE 2 overflights of ground-based ionosonde stations and higher altitude Alouette/ISIS topside sounder encounters with equatorial bubbles indicate that depletions of several orders of magnitude in $N_e$ correspond to bottomside penetrations, whereas depletions of one order of magnitude (or less) in $N_e$ are most common in true encounters with topside bubbles.

**Author's Reply**

In the present study, the satellite altitude is high enough ($\gtrsim 600$ km) so that bottomside penetration should not be a problem.

**F.Sluijter**

Can you comment on the physics of the stability of the tubes?

**Author's Reply**

If you mean by "stability" the lifetime of the bubbles with $n$ (inside) $< n$ (outside), it is mainly determined by the long diffusion time of the electrons across the magnetic field. Thus, electrons from the ambient plasma can hardly enter the bubble during its rise time up into the topside ionosphere. It is thus not "filled up" during its rise time.

**D.Anderson**

It is known that there exists a longitude dependence on low-latitude bubble formation and that, in fact, bubbles are primarily seen in summer months in the Pacific sector rather than winter months. Have you separated your bubble observations according to longitude to see if this is true?

**Author's reply**

In fact, there are observations at different longitudes, but the author would have to take a closer look to see if bubbles were seen in the summer months in the Pacific.

**T.Croft**

You have said that the bubble heating mechanism that operates at dawn cannot operate at dusk. Would you please explain why.

**Author's Reply**

At dawn the downcoming illuminated region of the ionosphere meets the upgoing bubbles, whereas at dusk the upgoing illuminated region runs in front of the also upward moving bubbles. This makes the difference.

**G.Rostoker**

I can't see how the satellite data distinguish between a bubble (with upper and lower altitude limits) and a field-aligned duct of great length along a field line. Have you any observational evidence which favors the "bubble" concept over the field-aligned duct picture?

**Author's Reply**

Yes, there are at least two evidences. One is from the early radar observations by Woodman and Laitoz. They clearly saw density voids with an upper and a lower altitude boundary. The second evidence is from plasma simulation results obtained by Haerendel and co-workers. They clearly obtained "bubbles".

# SIMULATION OF THE POLAR CAP F REGION IONIZATION USING AN EXPERIMENTAL CONVECTION ELECTRIC FIELD

by

C.Taieb
CRPE-CNET, 3 avenue de la République
92131 Issy-les-Moulineaux
France

**ABSTRACT**

The electrons and ions concentration over the polar cap F region has been calculated with a code running on a vectorial computer. Four coupled continuity equations, one for each ion specie, are solved together with the momentum equation, assuming that the ion velocities and temperatures are the same for all species. The transportation of the ionization is due to neutral wind drag, the electric field and the ambipolar diffusion velocity which includes a stress tensor for taking account of the anisotropy created by the intense electric field encountered in the polar cap. The horizontal ion transportation is due to the planetary convection electric field which is deduced from 24 hours EISCAT observations during a quiet day. Because the observations are limited northward in latitude it is necessary to interpolate the equipotential lines providing reasonable physical assumption. Then the experimental convection electric field is replaced by a theoretical one. Results from the simulation of the ions concentration with the two convection models are compared. They are compared also with the EISCAT electron density measurements.

## DISCUSSION

**G.Rostoker**

The high latitude anomalously low densities (compared to model predictions) observed at 0400 and 0500 MLT - where do they lie with respect to the auroral oval? Is it possible that the anomalously low number densities are associated with polar cap (or boundary plasma sheet) regions, whereas the good agreement between model and data at lower latitudes is associated with the central plasma sheet (i.e., equatorward portion of the auroral oval)?

**Author's Reply**

The low level of density at 0400 is associated with the high velocity observed by the radar. Whether this region is associated with the inside boundary of the auroral oval at that time or with the outside boundary, I don't know. I have not checked.

# AVERAGE DIURNAL AND SEASONAL PROFILES OF IONOSPHERIC STORMS

**G L WRENN**
Space Department, Royal Aerospace Establishment,
Farnborough, Hants GU14 6TD.

**A.S.RODGER**
British Antarctic Survey,
Madingley Road, Cambridge, CB3 0ET.

**H. RISHBETH**
Department of Physics, University of Southampton,
Southampton, SO9 5NH.

## SUMMARY

The reduction of Maximum Usable Frequency (MUF) during a geomagnetic storm represents a serious hindrance to the efficient operation of HF transmission links. This paper discusses possible improvements in forecasting services. Statistical analyses of ionosonde data from Slough, U.K. (52°N) are used to define patterns for the main phase effects of ionospheric storms and these are compared with similar patterns derived for other midlatitude stations. The storm induced changes in foF2 are primarily dependent upon local time, season and the strength of the geomagnetic disturbance, as quantified by the geomagnetic index ap(τ). The phase of the solar cycle controls average ionisation levels whilst short-term variations in solar activity have little effect. The computed mean patterns can be used to forecast foF2 for a few hours ahead, given the required geomagnetic input. Data from a complete solar cycle, 1371 through 1981, are used to determine the errors in the forecasts and to demonstrate that a useful advantage can be attained by this method. The root mean square error in foF2 for 89139 samples is 14.4% which compares favourably with the 17.0% that is obtained using forecasts based upon either quiet-time values or the previous day's measurements. Studies of error distributions highlight various limitations in the technique and indicate avenues to further improvement.

## 1. INTRODUCTION

Since the Maximum Usable Frequency (MUF) for radio communication between two widely spaced users is mainly controlled by critical frequency foF2, the 'negative storm effect' which considerably reduces foF2 can often prevent reception. The requirement to transmit and receive near to MUF is fostered by the need to escape from the congestion common to lower frequencies and the desire for greater bandwidth. Radio amateurs, particularly DX enthusiasts, are vulnerable because of frequency restrictions and relatively low signal strengths. The problem would be greatly relieved if good forecasts were available, but the long standing lack of understanding of ionospheric storms has continually thwarted this approach. Wrenn et al. [1] identified underlying patterns in the variability of foF2 observed over the South Atlantic\Antarctic region during ionospheric storms. Wrenn and Rodger [2] have since used Argentine Islands data to show how such patterns might be used to improve the forecasting of foF2 at midlatitudes.

The traditional concept of an ionospheric storm demands at least a 'start' and preferably an 'end'; unfortunately discrete and isolated storms are very rare. It is often difficult to identify defining characteristics although their complexity can sometimes be contained by tying certain features to *initial*, *main* and *recovery* phases. However, such analysis implies a dependence upon the elapsed time from the start (*storm-time*) and it is clear that the thermospheric response will then be a function of local-time at onset and the pre-storm conditions. Paper [1] avoided many of these problems by postulating that the ratio of foF2 to its quiet time value can be directly correlated to an integrated geomagnetic activity index, sp(τ). While this simple approach can only go part way to describing reality, it provides a basis for empirical modelling and offers a vital element of a prediction scheme which promises a significant advance on anything currently available. Two key components are involved, the determination of dependable quiet-time reference values and the development of solar and geomagnetic indices that reflect present and past activity with proper temporal weighting.

## 2. DATA ANALYSIS

The analysis described in [1] was developed intuitively, the steps being punctuated by critical surveys of a large quantity of data. The results seemed to justify the means but, in order to produce quantitative models and estimate appropriate confidence limits, it was necessary to test the various assumptions and simplifications used; this was done [2] with foF2 measured at Argentine Islands (65°S, 64°W). Here the adopted procedures are applied to data from Slough (51.5°N, 0.6°W) where ionospheric soundings have been made since 1932 and the first observations of the negative storm effect were made in 1935 [3]. The input data supplied by WDC-C1 are the hourly scalings of foF2 for a complete solar cycle, 1971-1981; the 3-hourly ap, and the daily 10.7 cm solar flux values (burst corrected) for the same period.

A key to describing the foF2 variability has been the introduction of a time-weighted accumulation of ap [1,4] and the selection of a suitable persistence factor $\tau$. $a_p(\tau)$ is a normalised summation of precessive ap values each multiplied by $\tau$ raised to a power equalling the number of 3-hour intervals which have elapsed; the persistence is defined as $-3/\ln \tau$ hours. $\tau$ = 0.75 or 0.8, persistence 12 to 15 hours, has been used as an appropriate value for the midlatitude ionosphere (but see section 2.3).

### *2.1 Quiet-time reference frequencies*

Over the eleven years, $a_p(0.8)$ is less than 4.7 for 10% of the 3-hour intervals. Using this as the criterion for 'quiet' conditions, 9188 foF2 values were extracted, about 30 for each hour of each month. To establish their dependence upon solar activity, the 10.7 cm radio flux was selected as being the most convenient and effective monitor. Daily values ranged between 65 and 290 ($\times 10^{-22}$ W m$^{-2}$) through the period but fig.1 shows that 27 day variations were large near solar maximum. Since these short-term changes did not modify foF2 significantly, see below, a smoothed index was required. As an 81 day running mean, recommended to users of MSIS [5], is only available retrospectively, this analysis uses $F_{10.7}(0.98)$ which is defined in the same way as $a_p(\tau)$. Since F10.7 is a daily monitor, $\tau$=0.98 corresponds to a persistence of $-1/\ln 0.98 \approx 50$ days. 0.98 was chosen to smooth out the short-term variations and match the amplitudes to those of the 81 day means but no formal optimisation was attempted. These features are clear from fig. 1. While $F_{10.7}(0.98)$ lags in phase near solar maximum, it does represent a reasonable compromise measure of solar activity - better than daily values, or monthly means which have large discontinuities. Plotting foF2, for particular months and hours, against $F_{10.7}(0.98)$ showed marked departures from linearity (see fig. 3) and it was decided to employ look-up tables and interpolation rather than curve fitting. Examination of the frequency distribution for the quiet days led to the selection of the following ranges for averaging:-

| | mean | s.d. | No of hours |
|---|---|---|---|
| 71 to 84. | 76 | 4 | 2960 |
| 85 to 112. | 99 | 9 | 2972 |
| 113 to 160. | 130 | 12 | 2826 |
| 161 to 221. | 148 | 12 | 2363 |

The quiet-time foF2 s were binned in terms of month, hour and range of $F_{10.7}(0.98)$ and then averaged to give a complete matrix. The lowest (solar minimum) and highest (solar maximum) frequency sets are plotted as fig. 2 with lines linking the 24 hourly means. These, together with the data points for the representative hours plotted as fig. 3, illustrate the typical accuracy that can be expected from linear interpolation between the four mean values utilised in the quiet-time model. These four samples, covering a range of season and local time, reflect the extremes in gradient and non-linearity; the scatter of points is encouragingly small.

For any hour of any day of any month it is now possible, with the known or estimated value of $F_{10.7}(0.98)$, to calculate the appropriate quiet-time foF2 (fq). This involves linear interpolation between the four activity levels and also between consecutive months, the averages being assumed to apply for the 16th of each month. Since $F_{10.7}(0.98)$ changes slowly, a reasonable estimate of fq can be made for several weeks in advance and longer term predictions are possible.

In order to check that short-term variations in the 10.7 cm flux really can be neglected, all 9188 'very quiet' foF2 s were compared with the corresponding model calculated values, the ratio (fo/fq) being plotted against the difference between the daily and accumulated fluxes [$F_{10.7} - F_{10.7}(0.98)$]. Fig. 4 shows no obvious trend within the scatter for which the best fit straight line is determined to be :-

$$fo = fq \{ 1.010 + 0.0010 [ F_{10.7} - F_{10.7}(0.98)] \} \quad \ldots (1)$$

This confirms that any additional dependence upon the *daily* $F_{10.7}$ is very weak.

### *2.2 Seasonal Patterns*

Following the treatment described in [1] the computed values of ln(N/No) for Slough, 1971-1981, were binned in station time (ST), month and range of $a_p(0.75)$. Fig. 5 shows the resulting seasonal trends in the diurnal profiles of the mean values for very disturbed conditions i.e. $a_p(0.75) > 30$. The smallest values of ln(N/No) occur in the early morning hours during March and April and, to a lesser extent, August and September; the only positive values are post noon in January. While the sharp features (e.g. 18UT, December) coincide with inadequate sampling, the essential profile, changing smoothly from month to month, conclusively establishes that storm modification of the F2-region is strongly time dependent. The storm effects at Slough are considerably smaller than those at Argentine Islands at the same geomagnetic latitude (54°) in the southern hemisphere, fig. 8 of [1]. The inference is that the local time dependence is influenced by 27.4° *geographic* latitude, perhaps sunrise and sunset times; the difference in longitude is less likely to be important.

### 2.3 Optimising the Persistence Factor

In [1] $\tau$=0.75 (12 hour persistence) was adopted simply by inspection but no real attempt was made to justify this. Diurnal profiles of ln(N/No) presented in figs. 5 & 6 of [1] show that increasing $a_p(0.75)$ corresponds with an enhancement of main phase storm effects and establishes that such effects are greatest during the summer months. The

coarse binning led to average values with large standard deviations and begged the question of a meaningful correlation between $\ln(N/N_0)$ and $ap(\tau)$. The optimum value of $\tau$ was found to be 0.8 for Argentine Islands [2] by selecting an interval exhibiting strong geomagnetic activity dependency and determining the maximum correlation coefficient for varying $\tau$. Results for other intervals and other stations are presented as fig. 6 and are interesting because they provide a measure of the times of thermospheric response to geomagnetic disturbance: 9 h at Halley Bay, 12 h at Port Stanley, 14 h at Argentine Islands and 30 h at Slough. A shorter time (~7 h) applies to the positive effect at Argentine Islands on winter evenings, but here the significance is much lower. The correlation coefficient curve peaks are not sharp and although $\tau$=0.8 is here used to order the data, previous analyses using $\tau$=0.75 are still valid; a slighter higher value (~0.9) might be marginally better for Slough (persistence may actually change through the solar cycle).

### 2.4 *Linear Fitting of Geomagnetic Activity Modification*

For every value of foF2 during 1971-1981 the corresponding fq, the ratio fo/fq and ap(0.8) were determined as inputs for linear regression analyses of each month-hour set. Measurements made within 9 hours of an SSC were eliminated because they do not fit into the main phase patterns which have been identified [1]. Fig. 7 gives examples of scatter plots for fo/fq as a function of ap(0.8); the best-fit straight lines are:-

(a) March, hour 03: $fo/fq = 0.96 - 0.0055 \times ap(0.8)$ with $r = -0.417$
(b) January, hour 17: $fo/fq = 1.02 + 0.0042 \times ap(0.8)$ with $r = 0.228$

These examples, taken for gradient extremes, show the typical quality of fit for the dataset. For Argentine Islands data, the gradient ranges between -0.013 and +0.009 and the correlation coefficients ($r$) are correspondingly greater (-0.793 & 0.445) [2].

Intercepts and gradients of the best-fit lines are stored as 24x12 matrices so that, given a computed or estimated value of ap(0.8), it is now possible to calculate foF2. TABLE 1 lists ratios (100 fo/fq) for four geomagnetic activity levels given by ap(0.8) equal to 6 (quiet), 15 (normal), 25 (disturbed), and 45 (very disturbed). The very disturbed values effectively contain the same information as fig. 5 but some refinement has been achieved by revising the bin boundaries to eliminate undersampling.

## 3. foF2 FORKCASTING

Hf radiocommunication (~2 to 40 MHz) over distances of hundreds to several thousand kilometres is strongly dependent upon reflection from the ionospheric F2 layer but the latter is highly variable. For many years there have been a short-term forecasting services which attempt to predict the deviations on a timescale which is useful for operational *frequency management*. MUF's are a function of layer height and thickness as well as foF2, but variability of foF2 represents the major obstacle to successful forecasting. 24 hours seems to be the right order of lead time required for short-term forecasts although there is no doubt that a 6 hour service could be utilised if it offered increased reliability; longer term predictions of future trends also have an important role. Empirical modelling of foF2 from a single ionosonde station is described above and the method can obviously be applied to other stations. The incorporation of the diurnal and seasonal patterns, described in relation to solar and geomagnetic activity indices, into forecasting algorithms should reduce the impact of variability. The prediction of quiet-time values is simple and can be carried out with confidence for *several weeks in advance*; solar cycle trends can also be established with reasonable accuracy. The inclusion of the geomagnetic storm effects is more difficult and depends upon the availability of ap. The introduction of digital magnetometers and data networks has made the rapid determination of ap (Kp) feasible and the Space Environment Services Center at Boulder provides continual updates of the latest values and 27 day predictions. Magnetic parameters, from which ap can be found, have recently become available in real time *'GIFTS'* (W.F.Stuart, private communication).

### 3.1 *A Test of the foF2 Forecasts.*

*It is difficult to establish how good a forecast must be for it to be useful, or how bad it can be before it becomes counter-productive, but the first step is to define a quantitative measure of its success.* It is easy to compare the forecast with subsequent measurement and determine the error. For a large number (N) of forecasts (ff) it is logical to compute a root mean square error as

$$\epsilon = \sqrt{\Sigma[(ff-fo)^2] / N} \qquad \ldots . (2)$$

which can be related to the corresponding mean frequency ($\Sigma$ fo)/N. The clear objective of improved forecasting must then be to minimise $\epsilon$. For a 24 hour forecast it is possible to use today's *measurement* as a forecast for tomorrow so any worthwhile scheme should do better than this. Using all the 89139 scalings during the eleven years, TABLE 2 lists mean foF2 for each hour set and compares the results of three forecasts. Adopting the previous day's measurement (A) gives rms errors between 0.70 and 1.25 MHz through the day, i.e. between 14% and 19% of the mean values. Adopting instead the quiet-time values (B) produces a different distribution with errors between 0.79 and 1.23 MHz; near midday there is little change but at all other local times the quiet-time assumption is worse. However, using the ap(0.8) dependence (C) does succeed in reducing the errors at every hour, ranging between 0.57 MHz at 05 h to 1.04 MHz at 11 h. It should be noted that the

$F_{10.7}(0.98)$ and $a_p(0.8)$ values used in (B) and (C) were those corresponding to the day and hour concerned. Since the initial phase of an ionospheric storm is not covered by the model, 3064 measurements made within 9 hours of a reported SSC were eliminated from the treatment for (C).

Further analysis of the (C) results gives a breakdown of the accumulated errors into month and year. Rms errors for each month, January to December, (all hours) are
0.73 0.89 0.93 0.88 0.82 0.67 0.62 0.66 0.89 0.97 0.91 0.83 MHz.
This pattern is interesting because it shows that forecasting is more difficult near an equinox in March or September\October than in the summer or winter. Rms errors for each year 1971 to 1981 (all hours) are
14.2% 15.4% 16.6% 15.9% 14.6% 14.6% 14.7% 15.7% 14.9% 14.4% 17.2%
of the corresponding mean frequencies. Whilst the latter increase towards solar maximum, the percentage errors are very similar and there is no obvious pattern: for every year the hourly distribution is similar.

## 4. DISCUSSION

The underlying causes of the time variations of both the height and plasma concentration at geomagnetic midlatitudes are still poorly understood. Empirical models will never achieve particularly high success rates. In parallel with practical forecasting attempts there is a need for more theoretical and modelling work to understand the fundamental physics of the problem; Rishbeth [6] describes the present state of knowledge.

Although the diurnal and seasonal patterns in foF2 can clearly be established, they are generally masked by considerable day-to-day variability which cannot be accounted for. This renders the process of forecasting unreliable and, whilst the treatment described here does offer some advance, it does not represent the breakthrough that is really needed. There are numerous limitations with the basic data and some of these have been highlighted by this analysis. The scaling of ionograms is something of an art and the problems associated with oblique echoes, F1 signal distortion, sporadic E, interference and dispersion can produce errors of several percent. When the inaccuracy of scaling was assessed to be greater than ± 2%, this is indicated by qualifying letters [7], which were ignored in this analysis because of their relative insignificance in the large data-base. Given the scaling resolution of 0.1 MHz for foF2 and the inclusion of some points which may be 20% high (E) or low (D), the overall accuracy is unlikely to be better than ~5%. The digitising of the data is also not without fault; a number of typing errors have been identified but other less obvious mistakes may remain undetected.

The 10.7 cm flux is probably a rather coarse monitor of the solar EUV fluxes which control the production of the F2-region [8] but it is difficult to anticipate any realistic alternative. Sunrise and, to a lesser extent, sunset cause some discontinuity and this leads to error because the analysis is limited by the hourly resolution of the foF2 measurements. For a particular hour this effect can be smoothed out within a monthly average.

The initial phase of ionospheric storms presents a problem because foF2 then behaves quite differently to the main phase patterns. Usually it exhibits a strong positive effect but the amplitude and duration of this feature depend upon the local time of the SSC (sudden storm commencement). While it might be possible to predict recurrent storms, this cannot be done accurately and no satisfactory model has yet been developed. At present one must accept that SSC's will come without warning and the best that can be done is to issue a warning and wait for the recovery. There is some compensation for communicators in that the negative effects and much lowered MUF's are usually delayed by several hours from the onset time.

The introduction of $a_p(0.8)$ has permitted the regression analysis which is the basis of the forecasting strategy described here. However, this derived index is far from ideal and its limitations [4] must not be ignored. For any ionosonde station it is likely that a local 'a' index would form a better starting point but such data are not generally available in a suitable format. It has been demonstrated that $a_p(0.8)$ serves the purpose better than $a_p$; Dst was tried but the results were poor, the persistence of the ring current is different to that of the thermosphere. If the ionospheric storm has part of its origin in the auroral region, then it makes sense to search for an foF2 dependence upon an index of the AE type; this has yet to be attempted. Scatter plots such as fig. 7 certainly show that the adopted linear functions of $a_p(0.8)$ are only justified by the poorness rather than the goodness of the fit, that is to say that any more sophisticated fitting is not warranted by the extent of the scatter; for Slough the latter would be slightly reduced by utilising $a_p(0.9)$. There is obviously room for improvement in the present treatment.

Since the data best fits a persistence of some 30 hours, it is clear that a forecast of the type proposed here is limited to a few hours but advances in the techniques for geomagnetic monitoring should permit the required index to be computed close to real time, or at least up to 6 hours previously.

The height of the F2 peak (hmF2) is important to the physics of the production and loss processes and the HF reflection; how this changes during an ionospheric storm is therefore a vital question. The apparent height h'F2 gives some information on layer thickness and underlying ionisation but a better estimate can be obtained from M(3000)F2

[9]. The digitised scalings of M(3000)F2 and foE are included in the data set and the calculated hmF2 could be treated in a similar way to foF2. The uncertainties will be greater but the broad-brush statistical approach should yield meaningful results. Initial inspection of 1973 and 1974 hmF2 values suggests that the storm induced height changes are small. The current deployment of digisondes with automatic scaling of hmF2 should provide important answers.

There is growing evidence for the effects of gravity waves on foF2. 10% variability on a time-scale of tens of minutes to an hour or two. The type of statistical treatment described here can in theory be extended over one or two dimensions with a number of stations but it will be difficult to achieve the necessary matching between the characteristic scale lengths and the sampling distances which are fixed and far from ideal.

Since ionospheric storms can last for several days and the pronounced negative effects are common, it seems inconceivable that the principal changes in F-region electron concentrations result from a variation in the local production rate of ionisation. Some enhanced particle precipitation at midlatitudes as an indirect consequence of substorm activity, or the decay of the ring current, might produce increased ionisation but this would appear to be of secondary importance, except perhaps during the initial phase of the storm. It is generally accepted that the loss rate modification, due to changes in atmospheric composition, is the most probable cause of observed NmF2 storm trends but it is difficult to develop a theory which accounts for the detailed features of the measurements. An enhancement of the $[N_2]/[O]$ ratio, generated by auroral heating, may be transferred equatorwards by large-scale circulation or gravity wave but sophisticated 3-D time dependent computer models of the global thermo- sphere currently cannot reproduce the effects seen at midlatitude. Neutral wind patterns are crucial [10] but these are not known with sufficient detail; their high variability near equinox may well account for the relatively poor performance of the above forecast at such times.

Although the storm effects at Slough are smaller than at Argentine Islands, it is possible to deduce average profiles which are meaningful - in spite of large day-to-day and storm-to-storm variability. They exhibit repeatable and consistent patterns in terms of the level of geomagnetic activity, season and solar cycle phase. Such patterns may well hold the key to the identification of the physical processes involved and point the way to the required advance in forecasting.

## 5. REFERENCES


1. Wrenn, G.L., A.S. Rodger and H. Rishbeth, Geomagnetic storms in the Antarctic F-region. I. Diurnal and seasonal patterns for main phase effects, *J.atmos.terr.Phys.* 49(9), 901-913, 1987.
2. Wrenn, G.L and A.S. Rodger, Geomagnetic storms in the Antarctic F-region. II. Towards a strategy for the improved forecasting of foF2, *Radio Science*, submitted March 1988.
3. Appleton, E.V. and L.J. Ingram, Magnetic storms and upper-atmosphere ionisation, *Nature*, 548, 1935.
4. Wrenn, G.L., Time-weighted accumulations ap(r) and Kp(r), *J.geophys.Res.* 92(9), 10125-10129, 1987.
5. Hedin, A.E., MSIS-86 thermospheric model, *J.geophys.Res.* 92, 4649-4662, 1987.
6. Rishbeth, H., T.J. Fuller-Rowell and A.S. Rodger, F-layer composition and thermospheric storms, *Physica Scripta*, 36, 327-336, 1987.
7. Piggott, W.R. and K. Rawer, *URSI Handbook of ionogram interpretation and reduction*, Elsevier, 1961.
8. Killeen, T.L., Energetics of the Earth's thermosphere, *Rev. Geophys.*, 25(3), 433-454, 1987.
9. Dudeney, J.R., The accuracy of simple methods of determining the height of the maximum electron concentration of the F2-layer from scaled ionospheric characteristics, *J.atmos.terrest.Phys.*, 45, 629-640, 1983.
10. Sojka, J.J., R.W. Schunk and G.L. Wrenn, A comparison of foF2 obtained from a Time-Dependent Ionospheric Model with Argentine Islands data for quiet conditions, *J.atmos.terrest.Phys.*, in press 1988.



## ACKNOWLEDGEMENTS

This work was initiated at the University of Southampton with the support of NERC grant GST.02.112. It has since been pursued with the encouragement Paul Cannon of Radio & Navigation Department at RAE; discussions with John Milsom of Marconi Research Centre have been helpful. Mike Hapgood of WDC-C1 at RAL supplied the data on magnetic tape and floppy disc.


Copyright © Controller HMSO, London, 1988

TABLE 1. fo/fq x 100 at disturbed times **SLOUGH**

| | | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Quiet ap(0.8)=6 | Jan | 99 | 98 | 100 | 98 | 97 | 94 | 92 | 97 | 98 | 98 | 102 | 103 | 102 | 100 | 101 | 100 | 101 | 104 | 105 | 104 | 100 | 100 | 99 | 100 |
| | Feb | 99 | 98 | 97 | 96 | 93 | 93 | 90 | 91 | 98 | 99 | 102 | 100 | 101 | 100 | 102 | 103 | 102 | 102 | 99 | 99 | 98 | 101 | 104 | 103 |
| | Mar | 99 | 96 | 95 | 96 | 93 | 95 | 98 | 95 | 94 | 95 | 99 | 99 | 98 | 99 | 100 | 101 | 100 | 98 | 98 | 104 | 100 | 97 | 99 | 100 |
| | Apr | 97 | 98 | 97 | 95 | 95 | 95 | 98 | 100 | 99 | 100 | 100 | 99 | 100 | 101 | 100 | 99 | 99 | 99 | 98 | 98 | 98 | 98 | 95 | 97 |
| | May | 97 | 96 | 96 | 96 | 94 | 93 | 93 | 92 | 93 | 94 | 96 | 97 | 97 | 97 | 98 | 99 | 97 | 99 | 99 | 99 | 96 | 94 | 96 | 96 |
| | Jun | 98 | 98 | 97 | 96 | 94 | 96 | 96 | 96 | 97 | 99 | 98 | 99 | 99 | 100 | 99 | 100 | 101 | 101 | 101 | 100 | 97 | 99 | 98 | 100 |
| | Jul | 101 | 101 | 102 | 100 | 99 | 97 | 96 | 97 | 96 | 98 | 97 | 98 | 99. | 00 | 103 | 102 | 103 | 99 | 102 | 99 | 99 | 98 | 99 | 101 |
| | Aug | 98 | 98 | 98 | 97 | 96 | 97 | 97 | 97 | 99 | 101 | 102 | 102 | 101 | 102 | 100 | 102 | 102 | 101 | 102 | 105 | 101 | 102 | 100 | 99 |
| | Sep | 99 | 100 | 97 | 97 | 96 | 97 | 95 | 93 | 92 | 91 | 91 | 95 | 94 | 95 | 97 | 97 | 100 | 95 | 96 | 101 | 99 | 100 | 98 | 100 |
| | Oct | 98 | 98 | 97 | 95 | 94 | 93 | 99 | 95 | 95 | 96 | 97 | 98 | 99 | 98 | 100 | 99 | 97 | 99 | 102 | 102 | 100 | 96 | 98 | 97 |
| | Nov | 101 | 101 | 101 | 101 | 97 | 93 | 93 | 95 | 96 | 97 | 98 | 99 | 100 | 99 | 102 | 101 | 99 | 101 | 102 | 101 | 100 | 102 | 104 | 104 |
| | Dec | 99 | 100 | 100 | 98 | 98 | 99 | 93 | 96 | 97 | 96 | 98 | 97 | 96 | 97 | 97 | 97 | 97 | 100 | 100 | 97 | 95 | 97 | 97 | 100 |
| Normal ap(0.8)=15 | Jan | 96 | 93 | 94 | 92 | 92 | 87 | 85 | 92 | 95 | 95 | 101 | 105 | 104 | 103 | 105 | 104 | 104 | 108 | 107 | 104 | 101 | 100 | 98 | 97 |
| | Feb | 95 | 93 | 91 | 90 | 87 | 86 | 85 | 87 | 92 | 92 | 96 | 95 | 97 | 97 | 99 | 100 | 99 | 99 | 98 | 97 | 94 | 97 | 101 | 99 |
| | Mar | 92 | 89 | 87 | 88 | 85 | 87 | 95 | 91 | 89 | 89 | 95 | 94 | 94 | 96 | 97 | 98 | 97 | 95 | 95 | 101 | 96 | 92 | 93 | 94 |
| | Apr | 90 | 90 | 89 | 87 | 88 | 90 | 93 | 95 | 94 | 95 | 96 | 94 | 95 | 97 | 96 | 95 | 95 | 95 | 95 | 95 | 94 | 92 | 89 | 91 |
| | May | 90 | 90 | 88 | 88 | 87 | 87 | 88 | 88 | 88 | 90 | 92 | 93 | 94 | 94 | 95 | 96 | 95 | 97 | 97 | 96 | 92 | 90 | 91 | 90 |
| | Jun | 92 | 92 | 90 | 88 | 88 | 90 | 90 | 90 | 91 | 92 | 93 | 94 | 94 | 95 | 96 | 97 | 98 | 97 | 97 | 96 | 93 | 94 | 94 | 95 |
| | Jul | 95 | 93 | 94 | 93 | 93 | 91 | 91 | 92 | 91 | 94 | 94 | 95 | 96 | 98 | 101 | 100 | 101 | 97 | 100 | 97 | 96 | 94 | 94 | 96 |
| | Aug | 92 | 92 | 90 | 90 | 89 | 91 | 92 | 93 | 96 | 98 | 100 | 100 | 100 | 101 | 99 | 100 | 101 | 100 | 101 | 103 | 98 | 97 | 95 | 93 |
| | Sep | 91 | 92 | 89 | 90 | 88 | 89 | 90 | 88 | 88 | 87 | 86 | 90 | 91 | 93 | 95 | 95 | 98 | 92 | 92 | 97 | 93 | 93 | 91 | 92 |
| | Oct | 93 | 93 | 91 | 89 | 87 | 87 | 95 | 90 | 89 | 90 | 91 | 94 | 95 | 95 | 97 | 96 | 94 | 97 | 100 | 98 | 96 | 92 | 94 | 93 |
| | Nov | 97 | 95 | 95 | 94 | 91 | 89 | 89 | 91 | 91 | 92 | 95 | 96 | 97 | 97 | 100 | 100 | 98 | 100 | 101 | 99 | 98 | 101 | 102 | 100 |
| | Dec | 95 | 95 | 93 | 91 | 90 | 91 | 85 | 90 | 93 | 93 | 96 | 97 | 97 | 99 | 98 | 97 | 98 | 102 | 99 | 95 | 94 | 97 | 97 | 97 |
| Disturbed ap(0.8)= 25 | Jan | 92 | 87 | 88 | 86 | 87 | 80 | 76 | 86 | 93 | 91 | 101 | 106 | 106 | 105 | 109 | 110 | 107 | 113 | 110 | 104 | 102 | 100 | 97 | 95 |
| | Feb | 90 | 87 | 85 | 83 | 79 | 79 | 80 | 83 | 86 | 85 | 90 | 89 | 92 | 94 | 96 | 96 | 96 | 96 | 96 | 95 | 90 | 92 | 97 | 95 |
| | Mar | 85 | 82 | 78 | 79 | 76 | 77 | 91 | 86 | 83 | 83 | 89 | 89 | 90 | 93 | 94 | 94 | 94 | 92 | 92 | 97 | 90 | 86 | 87 | 88 |
| | Apr | 83 | 82 | 80 | 78 | 79 | 84 | 88 | 89 | 88 | 90 | 90 | 89 | 91 | 92 | 91 | 91 | 91 | 91 | 91 | 92 | 88 | 86 | 83 | 84 |
| | May | 83 | 82 | 80 | 79 | 79 | 81 | 82 | 82 | 83 | 85 | 87 | 89 | 90 | 91 | 92 | 93 | 92 | 94 | 94 | 93 | 88 | 85 | 85 | 83 |
| | Jun | 85 | 85 | 83 | 80 | 81 | 83 | 83 | 84 | 84 | 84 | 87 | 89 | 89 | 90 | 92 | 93 | 95 | 92 | 93 | 91 | 89 | 89 | 90 | 89 |
| | Jul | 89 | 85 | 86 | 84 | 86 | 86 | 86 | 87 | 86 | 90 | 91 | 92 | 94 | 95 | 98 | 98 | 100 | 95 | 97 | 94 | 93 | 89 | 89 | 91 |
| | Aug | 85 | 84 | 82 | 82 | 82 | 85 | 87 | 88 | 92 | 95 | 98 | 98 | 98 | 99 | 97 | 98 | 99 | 99 | 100 | 101 | 94 | 92 | 90 | 88 |
| | Sep | 83 | 84 | 81 | 82 | 78 | 81 | 84 | 83 | 82 | 82 | 81 | 86 | 88 | 90 | 94 | 93 | 95 | 89 | 88 | 92 | 87 | 85 | 82 | 85 |
| | Oct | 89 | 87 | 84 | 83 | 79 | 81 | 90 | 85 | 83 | 83 | 84 | 88 | 91 | 92 | 93 | 92 | 90 | 94 | 97 | 93 | 91 | 88 | 90 | 88 |
| | Nov | 91 | 88 | 88 | 87 | 85 | 84 | 84 | 87 | 87 | 88 | 91 | 92 | 94 | 94 | 98 | 98 | 98 | 99 | 99 | 96 | 96 | 100 | 100 | 96 |
| | Dec | 91 | 89 | 85 | 83 | 82 | 83 | 77 | 83 | 89 | 90 | 95 | 96 | 97 | 101 | 100 | 98 | 100 | 104 | 97 | 93 | 93 | 96 | 96 | 94 |
| Very Disturbed ap(0.8)=45 | Jan | 86 | 79 | 78 | 76 | 78 | 69 | 63 | 78 | 88 | 86 | 100 | 108 | 110 | 109 | 115 | 117 | 111 | 119 | 113 | 104 | 104 | 100 | 96 | 90 |
| | Feb | 84 | 78 | 75 | 72 | 68 | 68 | 71 | 77 | 75 | 74 | 80 | 81 | 86 | 90 | 92 | 91 | 91 | 91 | 93 | 92 | 84 | 85 | 92 | 88 |
| | Mar | 74 | 70 | 65 | 66 | 63 | 64 | 85 | 78 | 74 | 74 | 81 | 82 | 85 | 88 | 90 | 90 | 89 | 87 | 87 | 92 | 82 | 77 | 77 | 78 |
| | Apr | 71 | 69 | 67 | 64 | 66 | 75 | 79 | 80 | 80 | 82 | 82 | 82 | 84 | 85 | 85 | 85 | 84 | 85 | 85 | 87 | 80 | 77 | 72 | 74 |
| | May | 72 | 71 | 67 | 66 | 68 | 72 | 73 | 74 | 75 | 77 | 80 | 83 | 85 | 86 | 87 | 89 | 88 | 90 | 90 | 89 | 81 | 77 | 76 | 72 |
| | Jun | 74 | 74 | 72 | 69 | 70 | 73 | 72 | 74 | 74 | 73 | 78 | 82 | 80 | 83 | 86 | 87 | 89 | 85 | 87 | 84 | 82 | 82 | 82 | 81 |
| | Jul | 78 | 73 | 74 | 72 | 76 | 76 | 78 | 79 | 79 | 85 | 85 | 87 | 90 | 92 | 94 | 95 | 98 | 92 | 94 | 91 | 88 | 82 | 82 | 82 |
| | Aug | 75 | 73 | 70 | 70 | 70 | 76 | 78 | 81 | 86 | 91 | 94 | 95 | 96 | 97 | 95 | 96 | 96 | 97 | 98 | 97 | 89 | 85 | 82 | 78 |
| | Sep | 71 | 72 | 69 | 70 | 64 | 69 | 75 | 75 | 74 | 74 | 74 | 78 | 82 | 86 | 92 | 89 | 92 | 84 | 82 | 84 | 78 | 74 | 70 | 72 |
| | Oct | 82 | 78 | 74 | 72 | 66 | 70 | 84 | 77 | 73 | 74 | 74 | 80 | 84 | 87 | 88 | 86 | 84 | 91 | 94 | 86 | 83 | 81 | 83 | 81 |
| | Nov | 84 | 77 | 78 | 76 | 75 | 76 | 77 | 80 | 79 | 81 | 85 | 88 | 90 | 91 | 95 | 95 | 96 | 97 | 97 | 93 | 94 | 99 | 96 | 90 |
| | Dec | 85 | 80 | 73 | 71 | 70 | 70 | 63 | 73 | 83 | 85 | 93 | 96 | 98 | 105 | 103 | 99 | 103 | 107 | 95 | 89 | 92 | 96 | 95 | 89 |
| | | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |

TABLE 2. FORECASTING ERRORS **SLOUGH**

Mean foF2's and rms Errors for the three methods
A. Using yesterday's value (88910 samples)
B. Using average quiet-time values (92203 samples)
C. Using linear functions of ap(0.8) (89139 samples)

| hr | foF2 | rms ERROR | | | hr | foF2 | rms ERROR | | |
|---|---|---|---|---|---|---|---|---|---|
| | MHz | A | B | C | | MHz | A | B | C |
| 0 | 4.4 | 0.77 | 0.82 | 0.65 | 12 | 7.7 | 1.19 | 1.13 | 1.00 |
| 1 | 4.2 | 0.74 | 0.82 | 0.62 | 13 | 7.7 | 1.14 | 1.07 | 0.98 |
| 2 | 4.0 | 0.73 | 0.84 | 0.60 | 14 | 7.6 | 1.09 | 1.03 | 0.96 |
| 3 | 3.8 | 0.72 | 0.85 | 0.60 | 15 | 7.5 | 1.03 | 0.98 | 0.91 |
| 4 | 3.6 | 0.68 | 0.85 | 0.59 | 16 | 7.2 | 1.03 | 0.97 | 0.93 |
| 5 | 3.7 | 0.64 | 0.79 | 0.57 | 17 | 6.8 | 1.04 | 0.98 | 0.93 |
| 6 | 4.1 | 0.70 | 0.83 | 0.60 | 18 | 6.4 | 1.04 | 0.95 | 0.91 |
| 7 | 4.9 | 0.84 | 0.94 | 0.69 | 19 | 6.1 | 1.06 | 0.97 | 0.91 |
| 8 | 6.0 | 1.04 | 1.11 | 0.81 | 20 | 5.6 | 0.97 | 0.97 | 0.87 |
| 9 | 6.9 | 1.18 | 1.21 | 0.93 | 21 | 5.2 | 0.92 | 0.96 | 0.84 |
| 10 | 7.4 | 1.25 | 1.23 | 1.01 | 22 | 4.8 | 0.85 | 0.89 | 0.77 |
| 11 | 7.7 | 1.25 | 1.19 | 1.04 | 23 | 4.6 | 0.81 | 0.84 | 0.73 |
| | | | | | all hours | 5.7 | 0.97 | 0.98 | 0.83 |

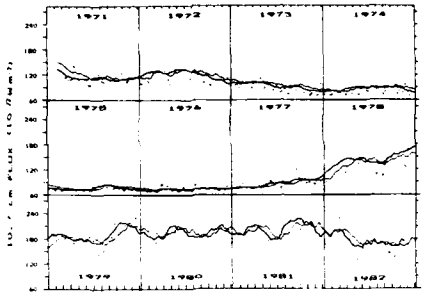


Fig.1. $F_{10.7}$ for 1971 to 1982:

daily values - points,
81 day running mean - thick line,
$F_{10.7(0.98)}$ - thin line.

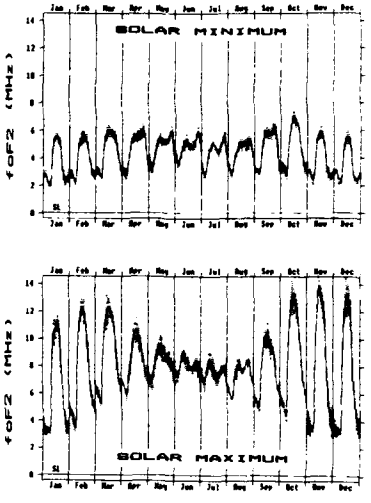



Fig.2. Slough quiet-time foF2 (MHz), monthly mean diurnal profiles. Hatching covers ± 1 s.d.

(a) solar minimum,
$F_{10.7(0.98)}$ = 76.4 + 3.9
(b) solar maximum,
$F_{10.7(0.98)}$ = 197.5 + 11.5

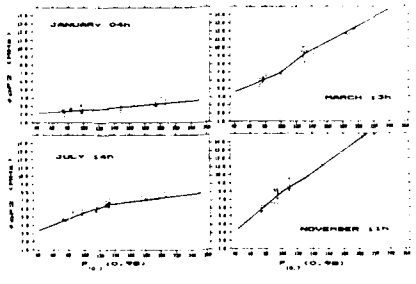


Fig.3. Quiet-time foF2 plotted against $F_{10.7(0.98)}$ with the mean model interpolations:

(a) January 04h (b) March 13h
(c) July 14h (d) November 11h

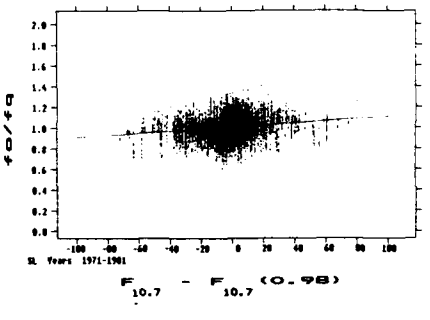



Fig.4. Quiet-time ratios of fo/fq versus *daily* $F_{10.7}$-$F_{10.7(0.98)}$

fq is the foF2 predicted with reference to $F_{10.7(0.98)}$

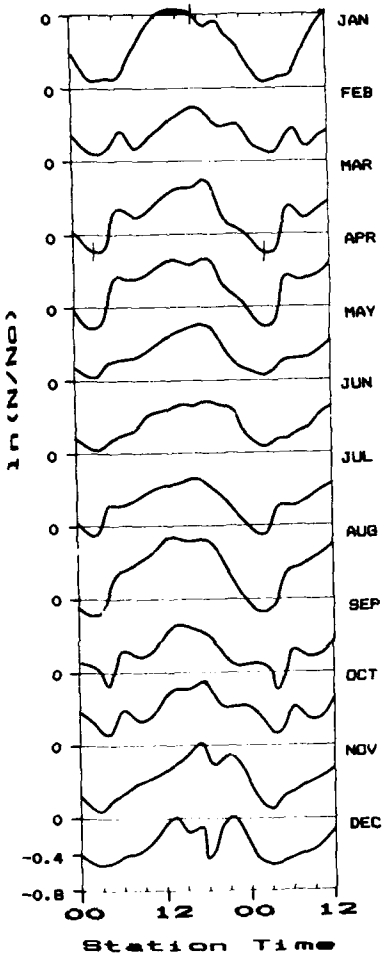



Fig.5. Variation of $ln(N/No)$ as a function of station time for each month for very disturbed geomagnetic conditions using data from Slough for the period 1971-1981.

Note that the first 12 hours have been repeated for clarity.

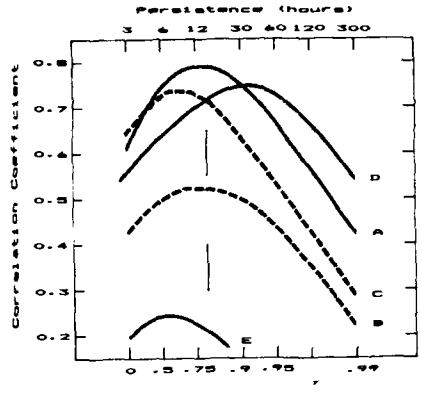



Fig.6. Correlation coefficients from linear fitting of $ln(N/No)$ to $a_p(\tau)$, as a function of $\tau$.

(A) Argentine Islands
( Dec\Jan\Feb, 2\3\4h LT )
(B) Port Stanley
( Nov\Dec\Jan, 5\6\7h LT )
(C) Halley Bay
( Nov\Dec\Jan, 2\3\4h LT )
(D) Slough
( Mar\Apr, 2\3\4h LT )
(E) Argentine Islands
( Jul\Aug, 19\20h LT )

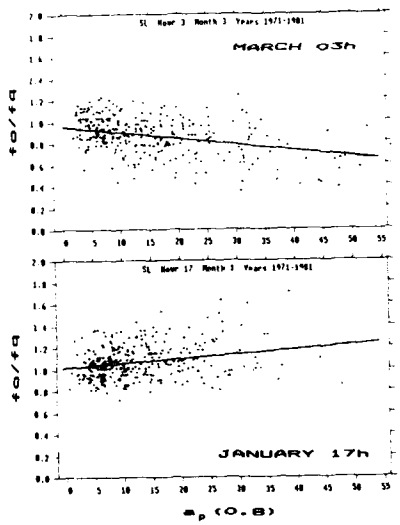



Fig.7. Scatter plots of fo/fq as a function of $a_p(0.8)$ with best fit lines.

(a) March 03h (b) January 17h

fq is the quiet-time foF2 interpolated with respect to day of month and F10.7(0.98)

## DISCUSSION

**K.Rawer**

With respect to conclusion #4 may I mention that analysis of extreme UV spectral measurements made by the AEROS satellites did not end-up with a fully satisfying description of the spectral changes in terms of *Ap* or $F_{10.7}$ We found that yet another parameter is needed which could only be obtained from the EUV-spectrum itself.

**Author's Reply**

I agree that the 10.7 cm monitor is a poor one for solar input to the thermosphere; there is evidence that it does not describe the EUV flux. Unfortunately, there is no better index available and the situation is unlikely to change in the foreseeable future, satellite measurements are few and far between.

**B.Reinisch**

You showed that using yesterday's values is not as good as your median method in forecasting. How did you use yesterday's values?

**Author's Reply**

Simply by taking the difference between 'yesterday' and 'today' as the error of the 'forecast'. The rms value of such 'errors' is about 20%. Any scheme of short term forecasting must do better than this. More sophisticated ways of using yesterday's values could be considered, but we have not attempted any.

# Medium Scale Structure of the F-region

Adolf K. Paul
Naval Ocean Systems Center, Code 542
Ocean and Atmospheric Sciences Division
San Diego, California,92152-5000
USA

SUMMARY

The standard picture of the F-region is based on hourly observations by ionosondes. More frequent recordings with advanced digital instruments show that very often F-region variations take place with periods of less than half an hour. This means that even if ionograms are taken in 15 minute intervals the actual variations of the ionosphere are still undersampled. Assuming that we observe the effects of propagating acoustic gravity waves, the undersampling of the oscillations results in an overestimation of the periods and the wavelengths. This in turn means that distances between existing and planned ionospheric stations are too large for correct spatial sampling of the waves. Therefore, at this time the observed temporal variations in combination with an assumed propagation velocity provide the only means to obtain first order information about the medium scale spatial structure of the F-region on a routine basis.

Digital ionograms recorded during 1980/81 at Brighton, Colorado show characteristics ( e.g. virtual height variations and doppler profiles) supporting the hypothesis that short term F-region variations are caused by acoustic gravity waves which seem to be present all the time with varying amplitudes. In addition, some direct evidence for the spatial structure can be obtained from the angle of arrival observations and their frequency or height dependence.

Comparisons of foF2 and MUF(3000) indicate that the main effect of gravity waves on the F-region structure is the variation of the height of the layer and to a lesser degree a variation of the maximum electron density.

## TEMPORAL AND SPATIAL STRUCTURE INFORMATION OBTAINABLE FROM VERTICAL SOUNDING

Modern technology now permits us to measure the radio phase of a single echo reflected from the ionosphere to one degree or better. It is this quantity with its high precision which provides much more detailed information about the ionosphere (Paul et al., 1976).

The comparison of the phases observed at three or more spaced antennas yields the angle of arrival of an echo which gives an estimate of the tilt of the surface of constant electron density at the reflection level. In many cases the angle of arrival measurement will underestimate the tilt angle at the reflection point if the angle between the electron density gradient and the vertical is a monotonic or almost monotonic function of the height. Such a situation exists, for example, during the undisturbed sunrise or sunset period, when the ray direction becomes almost horizontal in the vicinity of the reflection point for frequencies reflected close to the height of maximum electron density (Paul, 1985).

A Doppler frequency can be derived from the change of the phase with time. The Doppler velocity is then defined as the product of the radio wavelength and the Doppler frequency. It has to be mentioned that the observation of a Doppler frequency (or Doppler velocity) does not necessarily imply a motion of the reflecting area. The radio phase of an echo is proportional to the phase path which is the integral of the refractive index along the ray path. A decrease in the phase path, or a negative Doppler frequency, can be caused by a decrease in the distance to the reflection level or by a local increase in the electron density somewhere along the ray path. On the other hand, the combined effect of a local increase in the electron density and an increase in path length could result in no apparent change of the phase path, and hence zero Doppler frequency,but the group path would increase significantly in such a situation. Generally, the Doppler frequency (or the Doppler velocity) is an indicator for the presence of temporal changes in the ionosphere, but its interpretation, especially without comparison to the temporal changes of the virtual heights, is very difficult and often not unique (Bennet et al., 1986).

The change of the phase with frequency provides an improved estimate of the virtual path length, which is equal to the virtual height, if the propagation is strictly vertical. An accurate value for this quantity is highly desirable for a variety of analysis procedures like extrapolation of the echo trace for the determination of the critical frequency, electron density profile computation, studies of the temporal variations of the virtual heights at sets of fixed frequencies, etc.

ESTIMATES OF SPATIAL AND TEMPORAL STRUCTURE PARAMETERS

One of the most significant results obtained during the short period of operating a NOAA digital ionosonde at Brighton, Colorado was the detection of the high variability of the ionosphere, especially the F-region. The data reveal that the variations taking place are of a rather complex nature and that the magnitude of those oscillations can vary strongly with time. The MUF(3000), the maximum usable frequency over a 3000 km path, as derived from monostatic (quasi-vertical) ionograms, was selected as an indicator for the variability of the F-region. This quantity, or equivalently the propagation factor M(3000), is a standard propagation parameter routinely scaled at all ionospheric stations. By its definition, the MUF(3000) is also well defined and very sensitive to changes in the lower half of the F-region.

A typical example of MUF variations is shown in figure 1. Oscillations with periods mostly in the range from 20 to 30 minutes are clearly visible. There is, however, some indication that the variations may only be coherent over a few periods. Assuming that the observations shown in figure 1 are caused by propagating waves with a velocity of approximately 200 m/sec, the wavelengths for the above mentioned range of periods would be between 240 and 360 km.

Since the MUF(3000) is derived from the virtual heights at frequencies about 10% less than the penetration or critical frequency foF2, we expect to see similar variations of the virtual heights at fixed frequencies in this frequency range. An example of this behavior is shown in figure 2. The periods visible are again in the 20 to 30 minute range, but we also see that maxima and minima of the virtual heights appear later at lower frequencies (heights). The delay time is approximately 4 minutes over the frequency range shown, which corresponds to a true height range of approximately 50 km. While the first period shown in this figure may give the impression that the F-region as a whole may oscillate in height, the second period shows that the higher frequencies penetrate the layer for several minutes and are reflected again later, which can be interpreted as a local reduction of the electron density due to a temporary local expansion of the layer. Comparing the frequencies at the 600 km height level between the times 13:36 and 13:45 we find a temporary reduction of the maximum electron density of 4%.

A time lag between the variations of foF2 and the MUF(3000) is also clearly visible in figure 3. As mentioned earlier the critical frequency foF2 is determined by an extrapolation process based on the variation of the virtual heights with frequency over a small range close to the penetration frequency. This means that the effective height of variations visible in foF2 is lower than the height of the electron density maximum by approximately 10% to 15% of the half-thickness of the layer. On the other hand the height range which determines the MUF(3000) is on average 45% of a half-thickness below the peak. A height difference of 80 km between the two levels therefore seems to be a reasonably accurate estimate. The time lag corresponding to this height difference can be obtained by cross correlation of the two data sets. With the data shown in figure 3 a maximum of the cross correlation was found for a time lag of approximately 7.5 minutes. Combining the height difference and the time lag we obtain an apparent vertical downwards velocity component of 160 m/sec. Velocity estimates of this kind were computed for all high quality data sets which had the necessary temporal resolution of 20 or more ionograms per hour. The results were in the range 100 m/sec to 200 m/sec.

The changes taking place in the F-region can be rather complex, as is evident in figure 4. Here the apparent position of the reflection point is shown for a range of frequencies close to the penetration frequency, for three consecutive ionograms taken in three minute intervals. The numbers used as plotting symbols indicate the order in which the data were recorded. The lowest frequency used in all three traces was 11.3303 MHz, while the highest frequencies were 13.6662 MHz for the first and 13.7610 MHz for the second and the third ionogram. The low frequency end of each trace is always the point closest to the overhead point. We see that in the first ionogram the apparent echo position moves farther away from the overhead point in a north westerly direction with increasing frequency, but then suddenly changes direction by approximately 90 degrees toward the end of the trace. This change of direction occurs for the second ionogram at a lower frequency, and already starts at the lowest frequency used in this display in the third ionogram. With the electron density profile parameters computed from the ionograms we find that the point where the change of direction occurred dropped from a height of 344 km to a height of 305 km over the six minute time interval between the first and last ionogram, corresponding to an apparent velocity of 110 m/sec. The observations shown here have to be interpreted as a rapid change in height and time from one orientation of a prevailing strong F-region tilt to another one with the boundary between the two moving downwards. This example demonstrates that a time interval of three minutes can already be long in terms of F-region variations.

Some information about the spatial structure of the F-region can be deduced from the data shown in figure 5. The upper portion shows how the apparent reflection point moves away from the overhead point with increasing frequency during the sunrise period. A very similar trend is seen in the lower portion of this figure, but here the data were taken in the late morning. In both plots the points farthest away from the overhead point were observed at frequencies very close to the penetration frequency. For sunrise

conditions a model for the electron density distribution as a function of height and distance from the daylight boundary can be derived from the increase of foF2 as a function of time with reasonable accuracy. Ray tracing studies with such models show that the distance of the apparent reflection point from the overhead point d'(f) is to first order equal to the true distance of the reflection point from the overhead point (Paul, 1985). This can be explained by the facts that the virtual range overestimates the true distance to the reflection point and the observed angle of arrival underestimates its zenith angle with the two effects compensating each other to a large degree. The apparent distance d'(f) therefore can be used as a first order direct measure for a horizontal scale size - a distance over which significant changes of the electron density can be expected. More detailed information about local gradients or tilts can be obtained by comparison of d'(f) with the corresponding data of a sunrise electron density model.

Figure 6 gives an example of the variation of the apparent echo location over a longer period of time. Each data point shown was obtained from that portion of each ionogram where the MUF(3000) is determined, which means, as explained earlier, that the echoes came approximately from the middle of the lower half of the F-region. The data indicate that there is no preferred azimuthal direction for the angle of arrival or its equivalent, the horizontal component of the electron density gradient. The magnitude of the distance from the overhead point, however, seems to have a maximum in the north-west and south-east direction. It should be mentioned that this direction does not coincide with the magnetic declination, which is 14 degree east for this area.

A clear demonstration of wavelike variations in the F-region is given in figure 7. Here the Doppler frequency, as defined earlier, is shown for both magnetoionic components as a function of the radio frequency. The two traces display very similar sinusoidal variations of the Doppler frequency for approximately half a wavelength. The height difference over the corresponding plasma frequency interval, obtained from profile computation, is approximately 60 km, which implies a vertical wavelength of 120 km. Assuming a period of 20 minutes we obtain a vertical propagation velocity of 100 m/sec, which is in agreement with earlier findings.

## CONCLUSIONS

The examples reported here show that a variety of parameters indicating the presence of a medium scale structure of the F-region in space and time can be derived from single site observations with an advanced digital ionosonde. We have shown that the MUF(3000) can be observed with high accuracy and that it appears to be a reliable parameter describing the temporal variations of the lower half of the F-layer. Very frequently this parameter shows periodic behavior with periods mainly in the range from 20 to 30 minutes (Paul, 1988). Discrete events, as illustrated in figure 4, can occur within a much shorter time span and very significant changes can be observed in time intervals as short as 5 minutes. Time lags of similar magnitude are clearly detectable in the variations of the virtual heights at fixed frequencies or by comparison of two such MUF(3000).

The observation of spatial parameters is more limited than the measurement of temporal parameters. Angle of arrival measurements show the presence of tilts and their changes in time consistent with propagating wave phenomena. Estimates of the horizontal wavelengths, however, are only possible by making assumptions about the propagation velocity. Direct measurements of horizontal wavelengths would require observations from at least three sites separated by approximately half a wavelength. Estimates of the vertical wavelength are more reliable since in many cases they can be directly derived from Doppler profile data as shown in figure 7.

Unfortunately, very little can be learned about the magnitude and direction of the propagation velocity of the observed waves. The only directional information available are the estimates of the direction of the gradient of the electron density from the angle of arrival measurements. The direction of this gradient may be equal to the direction of propagation in some situations, e.g. during sunrise, but this may not be true in general. Similar remarks apply to the interpretation of the Doppler measurements which can only give a radial velocity component.

The results presented here give evidence that a medium scale structure is present in the F-region with a spatial scale of the order of tens of kilometers and a temporal scale in the order of minutes. Based on a very large sample of MUF(3000) data (Paul, 1988) from more than 20000 ionograms we have reasons to believe that this type of (moving) structure is present for almost all times, and only the magnitude of the variations is changing with time. Questions about the cause, origin and direction of propagation of such patterns remain still open, but might easily be answered by coordinated multi-static observations.

REFERENCES:

Bennet, J.A, and P.L. Dyson
The effect of small amplitude wave irregularities on radio wave observations of the ionosphere Radio Sci., 21, 375-387, 1986.

Paul, A.K., J.W. Wright and L.S. Fedor
The interpretation of ionospheric drift measurements - VI. Angle-of-arrival and group path (echolocation) measurements from digitized ionospheric soundings: the group path vector J.A.T.P., 36, 193-214, 1974.

Paul, Adolf K.
F-region tilts and ionogram analysis, Radio Sci., 20, 959-971, 1985.

Paul, Adolf K.
The MUF(3000) as an Indicator for F-region Variations, NOSC Technical Report 1204, January, 1988.

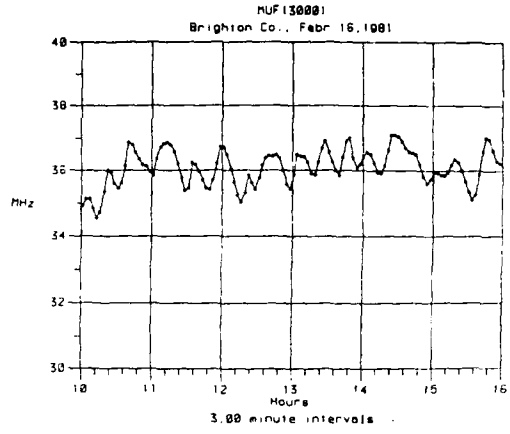



Figure 1. Temporal variation of the MUF(3000).

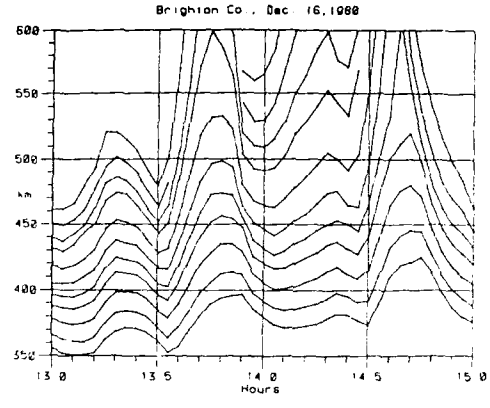



Figure 2. Virtual height variation at fixed frequencies.

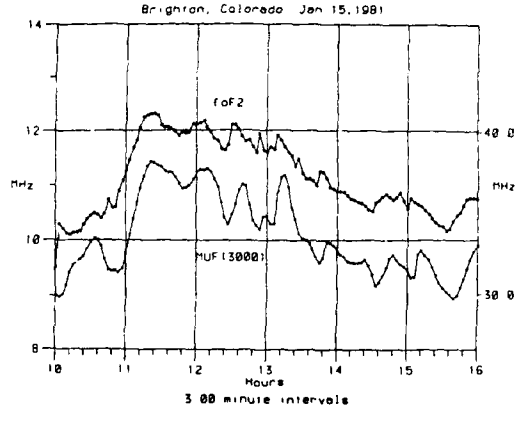



Figure 3. Comparison of foF2 with MUF(3000).

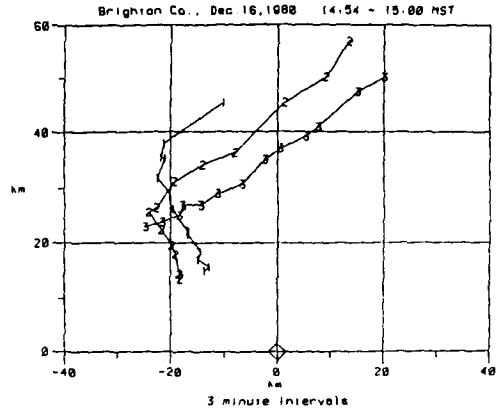



Figure 4. Apparent position of the reflection points for three consecutive ionograms.

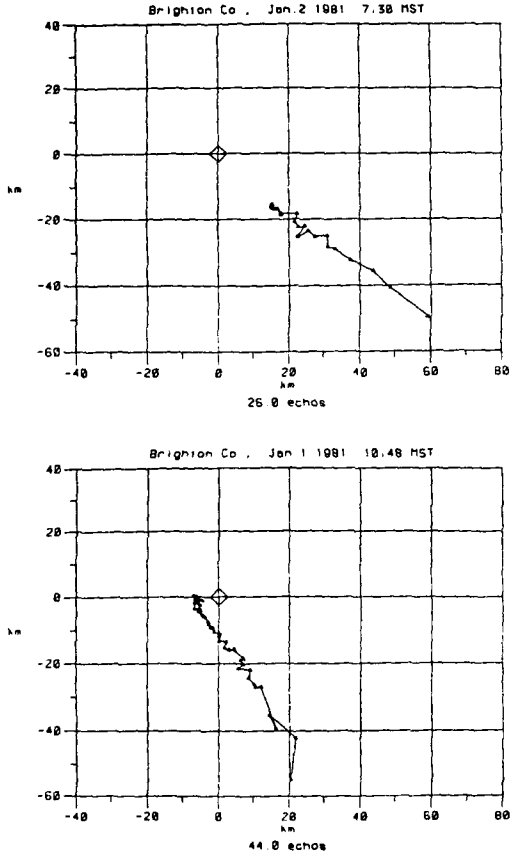



Figure 5. Change of apparent position of the reflection point with frequency during sunrise (top) and in late morning (bottom).

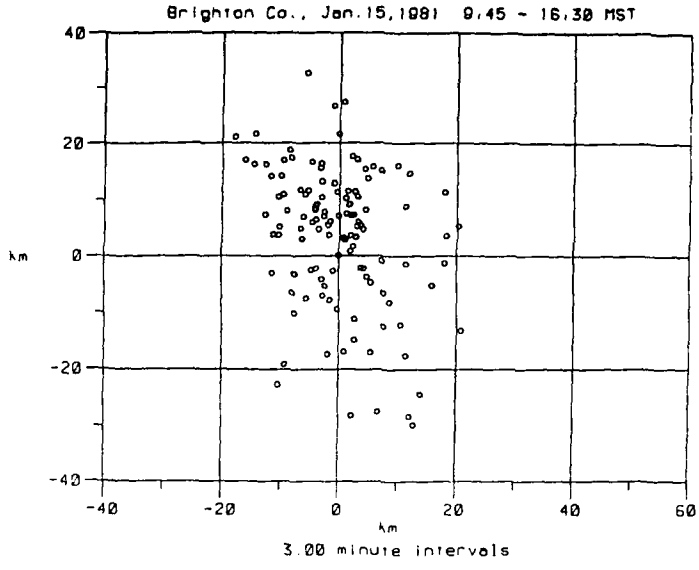



Figure 6. Horizontal distribution of apparent reflection points for echoes from the middle of the lower half of the F-region.

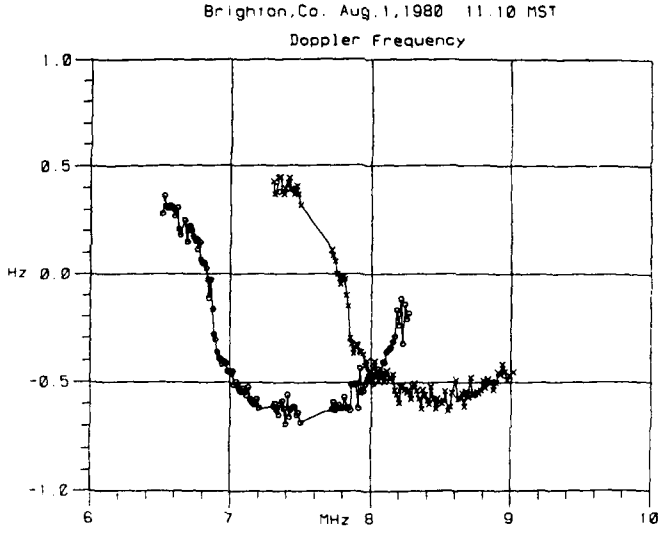



Figure 7. Variation of the Doppler frequency with radio frequency for both magnetoionic components.

## DISCUSSION

**B.Reinisch**

You showed convincing evidence of the fast variability of the ionospheric conditions requiring 5 minute sampling, or so. I agree with you and I am happy you pointed this out.

Questions:

1. You showed that the reflection point in the $F$ region moves by tens of kilometers within 8 seconds. Can you be certain that the $F$ layer was so perfectly stratified during the observation period so that only one reflection point exists at each moment. Our observations give evidence that this is rarely the case. How does the presence of several simultaneous reflection points affect the accuracy of your technique?

2. You interpreted the observed differences in Doppler shifts for the O and X traces as different vertical velocities at the O and X reflection heights. Since most echos are slightly oblique, is the Doppler not dominated by the horizontal drifts?

**Author's Reply**

1. Phases and amplitudes of all echoes are recorded as *four* antenna. Only if all four measurements are consistent with a plane wave the data are used for angle of arrival determination.

2. In our definition the Doppler velocity is equal to one-half of the temporal change of the phase path. For corresponding frequency pairs, there is only a small difference between the ordinary and extraordinary velocities which may be due to the different electron densities experienced along two different ray paths, but also due to the differences of the phase refractive indices for the two components.

K.C.Yeh

I have a two-part question.

(a) Is your computation of MUF (3000) applicable only to a spherically symmetric ionosphere?

(b) Your answer to part (a) seems to indicate that you have used a control point method, which ignores such effects as may be created by gravity waves. Is it therefore true that the real MUF (3000) may be more complicated than you have described? In general, you may even expect azimuthal asymmetry depending on the aspect of propagation path relative to the direction of gravity wave propagation.

**Author's Reply**

(a) The MUF computation is the equivalent of the transmission curve fitting.

(b) The quantity is used here only as an indicator for the $F$-layer variability, not for propagation prediction.

# VLF/LF RADIO WAVE STUDIES OF THE STRUCTURE AND VARIABILITY OF THE EQUATORIAL, MID-LATITUDE, AND POLAR IONOSPHERE BELOW 100 KM

P. A. Kossey, J. E. Rasmussen, and W. I. Klemetti
Ionospheric Physics Division
Air Force Geophysics Laboratory
Hanscom AFB, MA 01731, USA

## SUMMARY

This paper describes VLF/LF radio wave observations of structure and variability in the lower ionosphere. Emphasis is on data relating to the ionosphere below 100 km, obtained at equatorial, mid-latitude and polar locations, with a VLF/LF pulse ionosounder. The use of the data to help characterize the short- and long term variability of the lower ionosphere is considered, in addition to its use to infer specific structural features of the ionosphere, such as the C-layer below 70 km altitude. Data obtained within the polar cap during a number of solar proton events (SPEs) are also described.

## 1. INTRODUCTION

VLF/LF waves propagate to great distances in the waveguide formed by the surface of the earth and the ionosphere below about 100 km; i.e., the D-region during daylight and the lower E-region at night. Apart from the sunrise and sunset periods, VLF/LF propagation is considered to have stable phase and amplitude characteristics, especially when compared to those associated with the propagation of high frequency (HF) waves, which interact with the ionosphere over an altitude range that includes the D-, E- and F-regions. As such, VLF/LF forms the backbone of a number of long range navigation and emergency communication systems. In order to insure desired system performance it is necessary to know, or be able to predict with a high degree of certainty, the state of the regions of the ionosphere that will affect the given system. This implies a need for developing models of the ionosphere which can be used for propagation prediction purposes. These include phenomenological models, based on empirical data and theoretical models, derived from direct consideration of the physical processes which govern the structure and variability of the ionosphere. In general, a comprehensive theoretical model of the ionosphere is necessarily complex, owing to the large number and wide variety of physical processes that may have to be considered. In addition, the problem is made even more difficult in that some of the critical parameters that characterize those physical processes have not yet been adequately quantified. Nevertheless, numbers of empirical and theoretical models of the ionosphere have been developed to date, and that field of scientific endeavor remains an active one. Present activities are almost solely focussed on the region of the ionosphere above 100 km, due, in part, to the intensive use of the HF spectrum for civilian and military communication purposes, as well as the understanding that the physical processes in the E- and F-regions of the ionosphere are much more dynamic than those in the D-region. In order to encourage the expansion of such global modeling efforts to include more detailed descriptions of the lower ionosphere, this paper describes structure and variability in the ionosphere below 100 km, as evidenced by ground-based observations of VLF/LF radio waves.

## 2. VLF/LF PROBING OF THE LOWER IONOSPHERE

Variations in the propagation characteristics of VLF/LF waves are usually indicative of changes in the electron densities of the lower ionosphere. Under normal ionospheric conditions, these arise from a wide variety of physical processes, including, changes in the ionizing radiation from the sun, variations in the concentrations of key constituents in the atmosphere, changing solar zenith angle conditions, solar-cycle effects, and the varying influence of additional ionization mechanisms such as cosmic rays and meteor showers. Such effects produce short-term, diurnal, seasonal, latitudinal and solar-cycle variations in the propagation characteristics of VLF/LF waves. In addition, relatively large VLF/LF propagation effects accompany ionospheric disturbances, including those associated with magnetic storms, aurora, solar x-ray flares, high energy electron precipitation events, and solar particle events.

Descriptions and examples of VLF/LF propagation effects produced by the physical processes and phenomena outlined above can be found in a number of references, including Belrose (1982), Davies (1965), Watt (1967), and Dolukhanov (1971). For the most part, the VLF/LF observations of variability and structure in the lower ionosphere have been characterized in terms of amplitude and phase changes of continuous wave (CW) signals, propagating over long (megameters) paths. The technique has proven to be a very sensitive one to changes in the lower ionosphere and provides a relatively simple means to monitor very large geographical areas with only a few observation sites. The technique, however, suffers in that the observations tend to represent an average of the state of the lowest regions of the ionosphere over the length of the propagation path. Thus, VLF/LF propagation effects produced by relatively localized variations in the state of the lower ionosphere are integrated, or smoothed-out, making it very difficult to use the observations to quantify the ionospheric changes that produced them. This is usually true even in the case of severe ionospheric disturbances such as those accompanying large solar x-ray flares or solar particle (proton) precipitation events. Nevertheless, long path VLF/LF observations often cannot be explained using simple (usually exponential) electron density models of the lower ionosphere that have little or no structure.

Short path (steep-incidence) VLF/LF observations provide data on localized regions of the ionosphere, and exhibit features that more easily can be related to structural

features in the electron densities of the lower ionosphere. For example, short-path observations of 16 kHz signals were used by Bracewell and Bain (1952) to first suggest the presence of two ionized layers well below the ionospheric E-region. Subsequent research suggest that the lower of the two layers may be produced by cosmic rays, although the full characterization of this "C-layer" remains the subject of ongoing research and even some controversy, as discussed later. Phase and amplitude variations can be used to characterize steep-incidence VLF/LF observations of structure in the lower ionosphere. Of particular interest in this regard is the use of the data to determine effective plane wave reflection coefficients of the ionosphere (Bracewell et al., 1951, Lewis et al., 1973), since the manner in which the magnitudes of those coefficients vary, over a broad frequency range, can reveal significant structural features in the lower ionosphere. In addition, the experimental data can be compared directly with that obtained theoretically, using full-wave computational techniques in conjunction with electron density and collision frequency models of the ionosphere (Inoue and Horowitz, 1968). Thus, the VLF/LF reflection data can be used to develop phenomenological models of the lower ionosphere and to validate theoretical models, derived from consideration of the chemical and physical processes that occur in the upper atmosphere.

## 3. SHORT PULSE VLF/LF OBSERVATIONS OF THE LOWER IONOSPHERE

### 3.1. VLF/LF Pulse Ionosounding Technique

Figure 1 illustrates a high-resolution pulse ionosounder, developed by the U. S. Air Force Geophysics Laboratory, which has been used to obtain a variety of VLF/LF steep-incidence ionospheric reflection data. Details of the ionosounding technique are given in a number of papers describing its use in experiments conducted at low-, mid-, and high geomagnetic latitudes (Rasmussen et al., 1975; Lewis et al., op cit.; and, Kossey et al., 1983, respectively). The key feature of the technique is the transmission of VLF/LF pulses so short that, even at receiving locations a few hundred kilometers away, reflections from the ionosphere can be observed free of the ambiguities of ground wave and sky wave interference, which are characteristic of short-path CW measurements. The approach provides a relatively direct means of observing variations in the state of the lower ionosphere. The transmitted pulses are vertically polarized, but because of the effects of the geomagnetic field the reflected sky waves are elliptically polarized. Separate loop antennas are used to sense the "normal" ( $\parallel$ -) and "converted" ( $\perp$ -) components of the downcoming sky waves. Figure 1b gives typical examples of ground wave and sky wave pulses observed with the system, and Figure 1c shows the Fourier spectrum of a transmitted pulse to illustrate the wide range of VLF/LF frequencies that are available simultaneously to probe the lower ionosphere. As described by Lewis et al. (op. cit.), after Fourier analyses of the received pulses, the group delays between the sky waves and the ground wave can be used to determine effective heights of the ionosphere; and, the relative amplitudes of the sky wave and ground wave frequency components can be used to determine the normal ($_{\parallel}R_{\parallel}$) and converted ($_{\parallel}R_{\perp}$) reflection coefficients of the ionosphere (Bracewell et al., op cit.).

### 3.2. Observations of Short Term and Diurnal Variations in the Lower Ionosphere

Structure and variability in the lower ionosphere can readily be inferred from inspection of time-domain data. For example, the data of Figure 2 are typical of that observed under ambient conditions at mid-geomagnetic latitudes. Shown are ground wave and sky wave pulses for selected times throughout a single day, as observed over a 166 km path. In Figure 2a, observed 100 minutes before local noon, the ground wave occupies the first 140 $\mu$sec, and following a blank interval the sky wave onset occurs at about 160 $\mu$secs, corresponding, approximately, to an effective height of reflection of 67 km. Figure 2b shows a typical nighttime sky wave, beginning at about 240 $\mu$sec, corresponding to a reflection height of about 85 km. The nighttime amplitude of the sky wave is considerably larger than in the daytime.

Another nighttime example, observed 4 hours later (Figure 2c) shows a sky wave of almost twice the normal length, as if there were two reflections of nearly the same amplitude, one beginning at 240 $\mu$sec (from about 85 km) and the other at 320 $\mu$secs (from about 100 km). When successive pulse reflections are compared, an "extra" nighttime sky wave is sometimes seen to develop, descend, and merge with the normal sky wave in such a way as to temporarily cancel it out.

The data of Figure 2d, observed near sunset, shows a single sky wave that appears inverted when compared with the other examples, while the data of Figure 2e, observed near sunrise, shows an unusually weak sky wave.

Occasionally at night a very weak sky wave arrives from about the height of the daytime reflections. In the example of Figure 2f such a sky wave is seen arriving at approximately 155 $\mu$secs, corresponding to a reflection height of 65 km, while the main sky wave, beginning at 250 $\mu$secs (85 km reflection height) appears quite normal.

Figure 3 gives an example of a three-dimensional presentation of VLF/LF pulse reflection data that is useful for inferring changes in the electron densities and structure in the lower ionosphere. Shown are data acquired over a 246 km path in southeast Brazil (27 degrees, S latitude), consisting of waveforms stacked one behind the other in linear time progression from bottom-to-top for a 4-day period. For the display each waveform is a 15-minute average of approximately 325,000 pulses. The horizontal scale is linear in time (in microseconds), measured from the start of the ground wave. The waveforms represent the instantaneous wave amplitudes of the ground wave and normal ( $\parallel$ -) sky wave as a function of time.

The most obvious variations seen in the data of Figure 3 are those due to the changing solar illumination conditions. At night the ionospheric reflections are relatively strong, and come from an effective height of about 86 km. The lowering of the

effective height of the ionosphere to about 69 km at noon, in response to changing solar zenith angle conditions, is clearly seen in the data. In addition, the magnitudes of the daytime reflections are appreciably less than those at night, indicating substantial diurnal variations in the effective conductivity and/or structure in the lower ionosphere.

It has been possible to derive a phenomenological model of the electron densities in the lower ionosphere that produced the principal (86 km) VLF/LF pulse reflections observed at midnight in southeastern Brazil over more than a year's time in 1980-1981. The model is an exponential one, characterized by a $\beta$ = 0.8 km-1, h = 83 km, in the notation of Wait and Spies (1964). Theoretical values of reflection heights and plane wave reflection coefficients calculated with it agree very closely with those obtained experimentally across the 15-68 kHz band.

Closer inspection of the data (see circled areas) in Figure 3 reveals, however, structure in the nighttime ionosphere that cannot be accounted for by such a simple model. For example weak, moving, reflections can be seen coming from regions well above the 85 km altitude of the principal reflections. The reflections tend to begin at about 115 km and, in most cases, drift downward over a period of 6-8 hours before disappearing during the sunrise transition period. Sometimes, however, the reflections begin to drift upward again, about two hours before sunrise, as indicated in the Figure. Based on a very limited analysis of the data that has been done, the reflections appear to be best defined during the spring and fall months.

Detailed analysis of the daytime data also reveals VLF/LF reflection properties that cannot be accounted for by simple electron density models of the lower ionosphere that have no structure. Some evidence of such structure can be seen by close inspection of the daytime pulse reflections (e.g., within the square in Figure 3), which on certain days shows weak, relatively stationary, reflections that persist throughout the daylight hours followed by stronger reflections whose times-of-arrival vary in accordance with the changing solar zenith angle. The effects of ionospheric structure on the VLF/LF reflection properties of the lower ionosphere are more easily seen when the data are analyzed to determine effective plane wave reflection coefficients, as illustrated by data of Figure 4. Shown are noon reflection coefficients plotted over the 15-68 kHz band, derived from data obtained over a 215 km path in southeast Brazil. Data from only eleven days of observation are shown, but it is typical of that obtained over more than a year's time in 1980-1981. In general, those data were characterized by a relative minimum in the magnitudes of the reflection coefficients, as the frequency varied over the 15-68 kHz band. In the data of Figure 4, for example, the reflection coefficients decreased monitonically over the 15-50 kHz range, reached a minimum near 53 kHz, and then increased somewhat over the 53-66 kHz range. Experience with the VLF/LF sounding technique has shown that such a behavior is clear indication of interference phenomena produced by partial reflections from two layers or a sharp gradient of ionization in the lower ionosphere. Further, the frequency at which the minimum occurs can be used to estimate the difference in heights of the two layers that produced the interference pattern, when the geometry of the sounding experiment is taken into account (Kossey and Lewis, 1981). The spread in the data of Figure 4 indicates that, although the interference patterns were present throughout the period of the observations, the structure in the lower ionosphere that produced them varied somewhat from day-to-day.

It is interesting to note that when the noontime data acquired in Brazil during 1980-81 were restricted to below 50 kHz, an exponential model ($\beta$ = 0.43 km-1, h = 70 km) of the lower ionosphere was found that gave very good agreement between theoretical and experimental results. This indicates the importance of having as broad a range of frequencies as possible when employing VLF/LF techniques to probe the lower ionosphere.

The question arises as to what physical processes produce persistent structure in the lower daytime ionosphere. One, somewhat controversial, candidate is discussed below and involves the combined effects of cosmic rays and solar illumination on the lowest regions of the daytime ionosphere to produce a layer of ionization below the classical D region; the C layer.

### 3.3. VLF/LF Pulse Observations of the C Layer of the Lower Ionosphere

Figure 5 gives VLF/LF pulse reflection data which indicate more clearly the presence of a daytime reflecting layer below the classical D region. The data were acquired over a 263 km, mid-latitude path in winter, near a minimum in the solar cycle.

The nighttime portion of Figure 5a, from 0400 to 1200 UT and from 2200 to 0400 UT, show reflections beginning at about 160 $\mu$secs after the start of the ground wave, corresponding to effective reflection heights of about 82 km. The daytime portion of Figure 5a shows weaker reflections of a relatively complicated nature. To better understand them, Figure 5b shows a number of waveforms selected from Figure 5a in the period from 1130 to 2330 UT. Inspection of Figure 5b reveals that each daytime reflection consisted of two basic pulses that were slightly overlapping. The earlier pulse was nearly stationary throughout the daytime, but the later one had a varying delay, which caused the overlap to be greatest near local noon and least near sunrise and sunset. At 1230 UT, when the solar zenith angle at the mid-point of the path was about 91 degrees, the later pulse was especially weak, leaving only the earlier one; but, at 1330 UT, when the solar zenith angle was 82 degrees, the two pulses were merging. As described by Rasmussen et al. (1980), the effective heights of the later, higher altitude, pulse reflections seen throughout the daytime were in general agreement with a ln sec$\chi$ rule, following classical Chapman theory for the formation of the lower ionosphere.

It is believed that the ionization responsible for the earlier reflections was caused by cosmic rays and photodetachment, a mechanism proposed previously by other writers in connection with the C-layer of the ionosphere. In an analysis by Bain and Kossey (1987) a phenomenological model of the ionization that produced the early

reflections of Figure 5 was found consisting of a 6 km thick slab having a uniform conductivity of about 1.5 x 10-7 S/m, extending from 60 to 66 km altitude. With this conductivity and an assumed nominal collision frequency of about 2 x 10-7 s-1, the electron density associated with the layer is about 100 cm-3. However, it should be noted that this result was derived using data from a single day of observation, when the reflections from an apparent C-layer were especially strong and easily resolved from those associated with the classical D-region. Subsequent work by Rasmussen et al. (1982) indicates substantial day-to-day, seasonal, latitudinal and solar cycle variations in the nature of the daytime reflections observed with the VLF/LF ionosounder. For example, their data indicate that C-layer effects are weakest in the summer and strongest in the winter. In addition, the effects are very weak at low geomagnetic latitudes, in accordance with the expectation that the nearly horizontal geomagnetic deflects the incoming cosmic rays and prevents them from entering deeply into the earth's atmosphere. Finally, the data indicate that the cosmic ray effects are weakest near the maximum of the solar cycle, presumably due to the effects of the solar wind which are (inversely) strongest during that period.

Consideration of the effects of C-layers on ELF/VLF/LF propagation by Field and Lewinstein (1978) has produced some interesting results. They show calculated ELF/VLF/LF field strengths using daytime models of the lower ionosphere with, and without, a C-layer (given in Bain, 1974) to examine the effect of the layer on long range propagation. The model ionospheres used in their calculations are shown in Figure 6a, and the effects of the assumed layer on VLF/LF propagation for 20 kHz are illustrated in Figure 6b, and for 35 kHz in Figure 6c. The effects of the C layer on ELF propagation were very small, being less than 2 dB between the models, even out to distances in excess of 10 Mm. At 20 kHz the spread in calculated field strengths between the models was significantly larger, approaching a factor of four, or 12 dB, at great distances as shown in Figure 6b. At 35 kHz, the spread was considerably larger (Figure 6c), approaching a factor of almost forty, or 30 dB, at the longer ranges. For the calculations it was assumed that the C layer was present over the entire path, and hence, the results most likely represent extreme cases. However, the calculations indicate that the C layer acts as an absorber of energy in long wave propagation, and that those effects should be most noticeable on long path LF signals, rather than on ELF/VLF signals.

The characterization of the C layer and its effects on radio wave propagation remains a subject of ongoing research and even some controversy. In early theoretical work Nicolet (1958) proposed a low altitude layer of daytime ionization due to the combined action of cosmic rays and solar radiation. During the night the electrons produced by the cosmic rays rapidly attach to neutral particles to form negative ions which, because of the relatively large mass, do not respond significantly to VLF/LF radio waves. However, under solar illumination, electrons are liberated from the negative ions by photodetachment to produce an electron population which does interact with VLF/LF waves. This low altitude ionization is distinguished from that of the classical D region, which is under close control by solar zenith angle conditions, as discussed above in conjunction with Figure 5. Krasnushkin (1966) states that the C layer effectively controls the propagation of long radio waves, whereas Bremer and Singer (1977) found no need to invoke such a layer to explain low-frequency continuous wave propagation. On the other hand Aikin (1962), Hargreaves (1962) and Abdu et al. (1973) suggest that the C layer is only significant for short periods around sunrise. Bain and Harrison (1972) have derived a model of the lower ionosphere incorporating a distinct C layer for mid-latitude, summer noon, conditions during sunspot maximum periods. That layer is not greatly unlike the model derived from the data shown in Figure 5, which was for winter noon, sunspot minimum conditions. Finally, it should be noted that analytic models of the C layer, based on the known chemistry of the atmosphere (e.g., Molar, 1960, and Webber, 1962) show distributions of ionization that are too thick to account for most of the VLF/LF observations discussed above. Evidently, the global characterization of the structure and variability of the lowest regions of the daytime ionosphere remains an area requiring much further research.

### 3.4. Observations of the Structure and Variability of the Lower Polar Ionosphere

#### 3.4.1 Undisturbed Ionospheric Conditions

The state of the lower polar ionosphere, even under normal conditions, depends greatly on the unique solar illumination conditions that occur at high latitudes. As a result, it is necessary to describe its VLF/LF reflection properties for three different conditions: (1) day/night, corresponding to periods of the year when the sun is above and below the horizon for substantial portions of each day, such as during March and September; (2) night, corresponding to periods when the sun is always below the horizon, such as in December; and, (3) day, for periods when the sun is always above the horizon, such as in June. Typical VLF pulse reflection data received over a 106 km path in northern Greenland (Kossey et al., 1983) under those illumination conditions are given in Figure 7.

For example, in September, when the sun was above and below the horizon for about equal periods each day (Figure 7a), the sky waves were much stronger at midnight (0400 UT) than at local noon (1600 UT). Also, over the same period the pulses were reflected from effective heights that ranged from about 90 km at midnight down to about 78 km at noon. The variations closely followed the changing solar illumination conditions.

In December, when the sun was always below the horizon, very little structure or variability were seen in the data (Figure 7b). The sky waves were reflected from an effective height of about 86 km and remained relatively strong throughout each day.

In June (Figure 7c) the sun was continuously above the horizon. Nevertheless, the reflection heights varied from about 88 km at midnight down to about 78 km at noon. These diurnal variations indicate that the ionizing effects of the Lyman alpha radiation

responsible for the D region depended strongly on the solar zenith angle, even for angles considerably less than 90 degrees. Also, the June sky waves were considerably weaker than at any other time during the year. This suggests that there was an absorbing region of ionization well below the altitudes where the VLF pulses were reflected.

The midnight reflection heights in June, when the sun was always above the horizon, were actually higher than those for the months when the sun was below the horizon. Further, the midnight reflection heights in December were consistently lower than those at any other time of the year. It is believed that the reason for this is that a given electron production rate in the ionosphere occurs at a given pressure level, and that those pressure levels are at their lowest altitudes in winter (J. S. Belrose, personal communication, 1981).

Figures 8 and 9 summarize VLF reflection parameters derived from a large volume of data obtained over the same propagation path between 1976 and 1980. The data are indicative of substantial day-to-day, seasonal, and solar cycle variations in state of the lower polar ionosphere, under quiet, or undisturbed conditions. Figure 8 gives a compilation of data obtained at local midnight in 1976, near a minimum of the solar cycle, and in 1980, near a maximum of the solar cycle. Shown are averages of the normal reflection coefficients and effective heights of reflection over the 6-28 kHz band for the months of December, March, and June, and the standard deviations (or spreads) associated with the data. Figure 9 gives similar data, derived from noon data during the same periods. The corresponding data for the period 1977-1979 are not shown, but those data fit well within the bounds of the data shown in Figures 8 and 9.

The midnight reflection properties of the polar ionosphere in December and March (Figure 8) were very similar and did not change appreciably over the 1976-1980 period. In June, however, the standard deviations associated with the reflection data became much larger, as the solar activity increased from 1976 to 1980. Nevertheless, it is not certain from these data alone, that this difference in the data is indicative of a possible solar cycle effect, since the 1976 and 1980 data overlap to a great extent.

The corresponding data for noon (Figure 9) indicate little differences in the December reflection data over the 1976-1980 period. In March and June, however, there were significant variations. Specifically, the magnitudes of the reflection coefficients increased appreciably from 1976 to 1980, while the standard deviations decreased substantially. These features, along with the fact that the effective heights of reflection did not change appreciably, suggest that the lower polar ionosphere became more densely ionized and effectively more sharply bounded, as the solar activity increased.

Figure 10 gives (phenomenological) electron density profiles derived from the data summarized in Figures 8 and 9 (Kossey et al., 1983). The electron density model derived from the December data varies exponentially with altitude ($\beta$ = 0.65 km-1, h = 82 km). The model is appropriate for the 1976-1980 period. For the June data, however, separate electron density models were required for the 1976 and 1980 data, since those data indicate appreciable differences in the nature of the lower polar ionosphere as the solar activity increased. The model developed from the 1976 data (near solar minimum) exhibits a broad layer of relatively weak ionization in the 70-80 km altitude range, and has smaller electron densities than the model derived from the 1980 data, which varies exponentially with altitude ($\beta$ = 0.3 km-1, h = 73 km). Because of the very large day-to-day variations in the apparent state of the lower daytime ionosphere in March and September (near the equinoxes), it was not possible to derive simple ionospheric electron density profiles to characterize those periods. The large variations were due to the effects of the solar zenith angle conditions, which vary widely and change rapidly from day-to-day at high latitudes during the periods immediately preceding and following the equinoxes.

### 3.4.2 Disturbed Ionospheric Conditions

There were 29 long-lasting ionospheric disturbances observed over the 1976-1980 period in northern Greenland, caused by the precipitation of solar protons into the lower polar ionosphere following certain large solar x-ray flares. Although many of these solar particle events (SPE's) were not accompanied by significant increases in the absorption of high-frequency waves, all of them produced significant changes in the VLF reflection properties of the lower ionosphere. The effects were dynamic and often lasted for many days, indicating that the structure of the lower ionosphere varied considerably throughout those periods. As illustrated in Figure 11, the state of the lower ionosphere during energetic particle events depends greatly on the solar illumination conditions.

Figure 11a shows the effects of a very strong SPE (10-dB riometer absorption) that occurred in September 1978. Under quiet conditions (see lower portion of the data display) the sky waves at local midnight were much stronger than at local noon, and the effective heights of reflection varied from about 92 km to 82 km over the same period. The effects of the SPE are seen very dramatically in the data. Shortly after the onset, the effective height of the ionosphere dropped by more than 30 km, to an altitude below 50 km. The effects of the disturbance lasted for eight days, and there were considerable diurnal and day-to-day variations in the ionospheric reflections throughout that period.

The June data (Figure 11b) illustrate the nature of the VLF reflections for a SPE that occurred when the sun was continuously above the horizon. During quiet conditions the sky waves were relatively weak, and the heights of reflection varied from about 88 km at midnight down to about 75 km at noon. Even though the sun was continuously above the horizon, the sky waves had diurnal variations which closely followed the changing solar zenith angle. The effects of a SPE (6-dB riometer absorption) are also seen clearly in the data. The effective height of reflection was lowered by more than 15 km, to below 60 km, after which there were no discernible diurnal variations in the sky wave reflections until the effects of the disturbance were over, more than a week later.

The February data (Figure 11c) further illustrate the complexity of the interaction between energetic particles and solar radiation and the resultant VLF reflection properties of the disturbed polar ionosphere. During quiet conditions (see the early portion of the data) there were no significant diurnal variations in the sky waves, since the sun was essentially below the horizon all day. Shortly after the onset of a strong SPE (6-dB riometer absorption), the effective height of the ionosphere dropped by more than 28 km, and even at night during the first day of the event it remained below 60 km, because of a continuing high particle flux rate. For the next several days, however, there were large diurnal variations which were not present before the onset of the event. The variations were caused by a complex combination of processes, including the production of electrons by the effects of the incoming particles, the attachment of electrons to neutral molecules to form negative ions when the sun was well below the horizon, the freeing of attached electrons as the sun approached the horizon, and the recombination of electrons with positive ions to form neutral molecules. Generally, however, it appears that the diurnal variations were driven by attachment and photodetachment processes, rather than by photoionization by Lyman alpha, the radiation primarily responsible for the formation of the D region under quiet ionospheric conditions.

Figure 12 gives models of the ionosphere derived from the noon data of the September SPE shown in Figure 11a, obtained using a mathematical inversion technique, appropriate for isotropic propagation, developed by Warren et al. (1981). In Figure 12 the dashed curve gives an ionospheric profile representative of ambient daytime polar conditions (Reagan et al., 1981). It has relatively low electron densities (less than a few hundred electrons per cubic centimeter), even for altitudes well above 70 km. The profile derived from the VLF data obtained at noon on the first day of the SPE (September 23 profile) is dramatically different. It is very sharply bounded in the 48-50 km range and has a density of a few thousand electrons per cubic centimeter at 50 km. The profile derived for the second day of the event is similar, except that it is at slightly higher altitudes. The remaining profiles illustrate the slow recovery of the ionosphere toward quiet conditions, which occurred over a period of more than a week (Kossey et al., 1983). As illustrated above, the structure of the disturbed polar ionosphere below 100 km undergoes substantial changes over the course of a solar proton event, and each SPE has, in effect, its own characteristics depending on its severity and the solar illumination conditions that are present.

## 4. USE OF VLF/LF PROPAGATION DATA TO CHARACTERIZE THE LOWER IONOSPHERE

The inversion of steep-incidence VLF/LF reflection data, such as that discussed above, to obtain electron density models of the lower ionosphere is not an easy task. Under quiet ionospheric conditions it is especially difficult, since usually, the polarization rotation effects of the geomagnetic field cannot be ignored. Nevertheless, mathematical approaches that employ full-wave and iterative computational techniques have been developed and applied with some success (e.g., Shellman, 1970). Under disturbed conditions and certain daytime ambient conditions, when the VLF/LF reflections are controlled primarily by ionization below about 70 km, the geomagnetic field effects are greatly diminished and the ionosphere can be considered to be isotropic. Under such conditions the mathematical inversion problem becomes somewhat simpler. Warren et al. (op cit.) have developed an inversion technique, appropriate for isotropic propagation, which has been used to derive conductivity profiles of the severely disturbed polar ionosphere. Such inversion techniques require much care and knowledge in their application to determine models of the lower ionosphere; i.e., they cannot be routinely used. In addition, a problem with profiles calculated by mathematical inversion is that of nonuniqueness, which can be caused by either incompleteness of data or the nonlinear dependence of the reflected signal with the propagation medium. In addition, the profiles characterize narrow regions of the ionosphere, since the propagation data contain information about only those altitudes where the ionosphere interacts appreciably with the reflected wave. For VLF/LF that altitude range is generally less than about ten kilometers in extent.

The altitude constraints are even more severe if long path propagation data are used. The data in this case are the attenuation rates and phase velocities of propagating waveguide modes, and effectively, the analysis is a trial-and-error technique to find an ionospheric conductivity profile that provides a waveguide mode or wave hop structure agreeing with the observed distribution of radio field strength (Crain, 1970). Although such models may not be consistent in many respects with those derived from detailed analyses of the aeronomy of the upper atmosphere, they have found widespread application in long wave propagation prediction codes.

In this regard, the electron density models described by Morfitt et al. (1982), derived from analyses of a large volume and a wide variety of VLF/LF propagation data, have found widespread use for long path VLF/LF propagation prediction purposes. The models are very simple, varying exponentially with altitude. They vary according to day or night conditions, winter or summer conditions, and with latitude somewhat. The ionospheric models also depend on the frequency of the waves to be propagated, which clearly indicates their propagation, rather than geophysical, origins.

Although exponential models of the lower ionosphere are widely used to estimate VLF/LF propagation parameters, they often do not describe detailed features of VLF/LF experimental observations, except for certain, very specific propagation paths. Usually the ones from which the propagation data came which led to their development in the first case. Even then, interesting differences in predictions and experimental observations occur. For example, some of the models provide very good estimates of nighttime VLF/LF signal strengths, for waves propagating to the east, over water, at mid-latitudes. For propagation to the west, however, results using the same models are much poorer, indicating that propagation to the west may be more sensitive to fluctuations or

structure at higher altitudes in the nighttime ionosphere than is propagation to the east (Pappert and Hitney, 1987).

These and other ionospheric models commonly used for VLF/LF propagation prediction purposes do not incorporate any structural features, such as a C layer. Yet such a layer may greatly influence the propagation of radio waves in the higher VLF and LF band. Nor do the models account for the effects of a varying solar zenith angle, which may be large for long propagation paths, even under fully lit (daytime) conditions.

Similarly, the models do not provide adequate predictions of VLF/LF signals in the polar regions, which is not surprising given the high degree of variability associated with the VLF/LF reflection properties of the lower polar ionosphere, even under quiet conditions, as discussed above in conjunction with Figures 8 and 9. For disturbed polar conditions, such as SPE's, there are no dynamic models of the lower ionosphere that can be used to adequately predict the performance of VLF/LF systems. As discussed above, however, this too is not surprising since the structure of the lower polar ionosphere under disturbed conditions varies so widely from event to event, season to season, and at different times throughout the event.

Finally, there are other well known geophysical events which alter the structure of the lower ionosphere and significantly affect radio wave propagation, which are not accounted for in models of the lower ionosphere. Solar x-ray flares, for example, may occur many times during a single day during periods of high solar activity, and seriously degrade the performance of a variety of systems across the VLF/LF/HF/VHF bands. Yet the effects of such flares on the lower ionosphere have not yet been adequately characterized or incorporated into global, dynamic, models of the ionosphere.

## 5. CONCLUSIONS

In summary, there are many geophysical factors that produce variability and structure in the ionosphere. As a result of ongoing intensive research, including coordinated worldwide campaigns, many of these are becoming much better understood and quantified. For example, with the availability of solar x-ray and solar proton data in near real-time, from satellite monitors, the possibility exists for exploiting those capabilities to develop dynamic models of the disturbed lower ionosphere on a global basis. Similarly, data on other geophysical effects which produce appreciable changes in the structure of the lower ionosphere, including cosmic rays, high energy electron precipitation events, and even meteor showers should eventually be included in evolving global dynamic models of the lower ionosphere.

There continues to be fertile scientific research on characterizing the variety of physical processes that occur in the magnetosphere, ionosphere, and the atmosphere and their complex interactions which eventually govern the structure and variability of the electron and ion distributions in the ionosphere. For the most part these activities are directed to the development of models of the ionosphere above about 100 km, perhaps due to the wider use of communication and surveillance systems which are more directly affected by changes in the E- and F-regions, and to the general knowledge that the lower ionosphere is a relatively stable region. As discussed above, however, VLF/LF observations indicate that the ionosphere below 100 km exhibits substantial structure and variability, not only during ionospheric disturbances, but also under quiet conditions. Thus, as interest and progress grows in the development of global dynamic models of the ionosphere, it is hoped that the region below 100 km remains an important part of the definition of "ionosphere".

## 6. REFERENCES


Aarons, M.A., S. Ananthakrishnan, B.A. Krishnan, and S. Massambani (1973), Cosmic ray ionization in the D region at sunrise; Evidence from VLF measurements, Radio Sci. 8:733-736.

Aikin, A.C. (1962), The sunrise absorption effect observed at low frequencies, in Radio Wave Absorption in the Ionosphere, 287-300, Pergamon, New York.

Bain, W.C. (1974), The use of vlf propagation in ionospheric modelling, ELF-VLF Radio Wave Propagation, Proc. NATO Adv. Study Inst., D. Reidel Publishing Co.

Bain, W.C., and P.A. Kossey (1987), Characteristics of a reflecting layer below the classical D region, J. Geophys. Res. 92:12,443-12,444.

Belrose, J.S. (1982), The Propagation Medium: An Overview, AGARD Conf. Proc. 305:1-1-1-16.

Bracewell, R.N., K.G. Budden, J.A. Ratcliffe, T.W. Straker, and K. Weekes (1951), The ionospheric propagation of low and very low freque cy radio waves over distances less than 1000 km, Proc. IEE. 98:221-236.

Bracewell, R.N., and W.C. Bain (1952), An explanation of radio propagation at 16 kc/sec in terms of two layers below E layer, J. Atmos. Terr. Phys. 2:216-225.

Bremer, J., and W. Singer (1977), Diurnal, seasonal and solar-cycle variations of electron densities in the ionospheric D- and E-regions, J. Atmos. Terr. Phys. 39:25-34.

Crain, C.M. (1970), Ionospheric probing with long wavelength radio waves, J. Atmos. Terr. Phys. 32:551-566.

Davies, K. (1965), Ionospheric Radio Propagation, Nat. Bur. Stds. Monograph 80, U.S. Govt. Printing Office, Washington, D.C.

Dolukhanov, M. (1971), Propagation of Radio Waves, Mir Publishers, Moscow.

Field, E.C., and M. Lewinstein (1978), Analysis of electron density-profiles in the lower ionosphere, Rome Air Development Center Tech. Rep.:78-67, Griffiss AFB, New York, AD A057254.

Hargreaves, J.K. (1962), The behaviour of the lowest ionosphere near sunrise, J. Atmos. Terr. Phys. 24:1-7.

Inoue, Y., and S. Horowitz (1968), Numerical solution of full-wave equation with mode coupling, Radio Sci. 1(8): 957-970.

Kossey, P.A., and E.A. Lewis (1981), The reflection properties of conducting slabs, Rome Air Development Center Tech. Rep.:80-371, Griffiss AFB, New York, AD A098940.

Kossey, P.A., J.P. Turtle, R.P. Pagliarulo, W.I. Klemetti, and J.E. Rasmussen (1983), VLF reflection properties of the normal and disturbed polar ionosphere in northern Greenland, Radio Sci. 18:907-916.

Krasnushkin, P.E. (1966), Influence of the solar wind on the C-layer of the earth's ionosphere, Geomagn. Aeron. 6:452-453.

Lewis, E.A., J.E. Rasmussen, and P.A. Kossey (1973), Measurement of ionospheric reflectivity from 6 to 35 kHz, J. Geophys. Res. 78:3903-3912.

Molar, W.F. (1960), VLF propagation effects of a D region layer produced by cosmic rays, J. Geophys. Res. 65:1459-1468.

Morfitt, D.G., J.A. Ferguson, and F.P. Snyder (1982), Numerical modeling of the propagation medium at ELF/VLF/LF, AGARD Conf. Proc. 305:32-1-32-14.

Nicolet, M. (1958), Aeronomic conditions in the mesosphere and lower thermosphere, Sci. Rep. 102, Penn. State Univ. Ionos. Lab., State College, Pa.

Pappert, R.A., and L.R. Hitney (1987), Empirical modeling of nighttime easterly and westerly VLF propagation in the earth-ionosphere waveguide, Ionospheric Effects Symp. Proc. 87:63-74, Naval Research Laboratory, Washington, D.C.

Rasmussen, J.E., E.A. Lewis, P.A. Kossey, Reginaldo Dos Santos, Edson De Freitas Coutinho, R.C. Kahler, and W.I. Klemetti (1975), Low frequency wave-reflection properties of the equatorial ionosphere, Air Force Cambridge Res. Lab. Tech. Rep.: 75-0615, Hanscom AFB, Massachusetts, AD AC25111.

Rasmussen, J.E., P.A. Kossey, and E.A. Lewis (1980), Evidence of an [illegible] reflecting layer below the classical D region, J. Geophys. Res. [illegible].

Rasmussen, J.E., P.A. Kossey, and J.P. Turtle (1982), VLF/LF [illegible] below the ionospheric D-region, AGARD Conf. Proc. 305:9-1-9-[illegible].

Reagan, J.B., R.E. Meyerott, R.C. Gunton, W.L. Imhof, E.E. Gaines, and T.R. Larsen (1982), Modeling of the ambient and disturbed ionospheric media pertinent to ELF/VLF propagation, AGARD Conf. Proc. 305:33-1-33-10.

Shellman, C.H. (1970), Electron density distributions in the lower ionosphere with associated error limits derived from VLF and LF sounder data, Radio Sci. 5:1127-1135.

Wait, J.R., and K.R. Spies (1964), Characteristics of the earth-ionosphere waveguide for VLF radio waves, NBS Tech. Note 300, Inst. for Telecommun. Sci., Boulder, Colo.

Warren, R.E., E.C. Field, and C.R. Warber (1981), Calculation of ionospheric conductivity profiles by inverting VLF/LF reflection data, I, Isotropic propagation, Rome Air Development Center Tech. Rep.:81-286, Griffiss AFB, New York, AD A108577.

Watt, A.D. (1967), VLF Radio Engineering, Pergamon Press, New York.

Webber, W. (1962), The production of free electrons in the ionosphere D layer by solar and galactic cosmic rays and the resultant absorption of radio waves, J. Geophys. Res. 67:5091-5106.

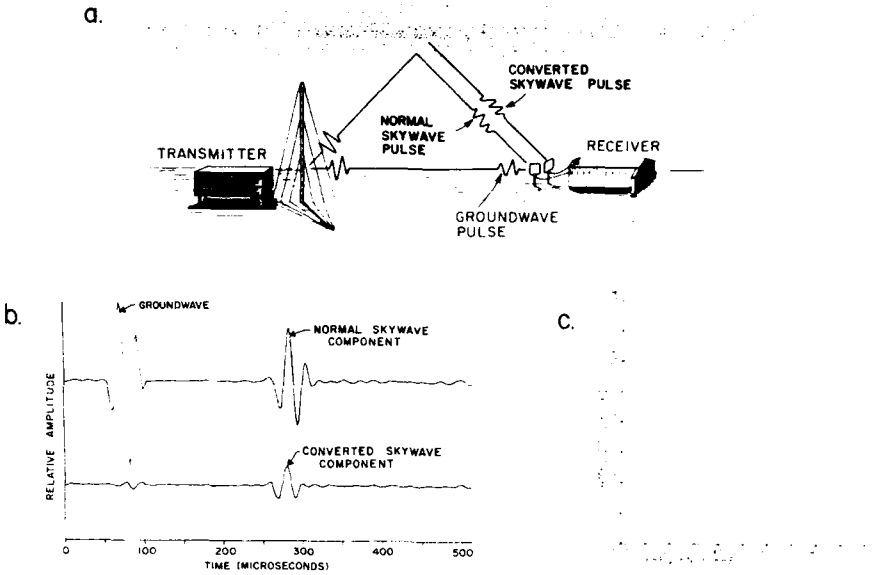



Figure 1. (a) Basic VLF/LF ionosounding technique.
(b) Example of observed waveforms.
(c) Spectrum of typical transmitted pulse.

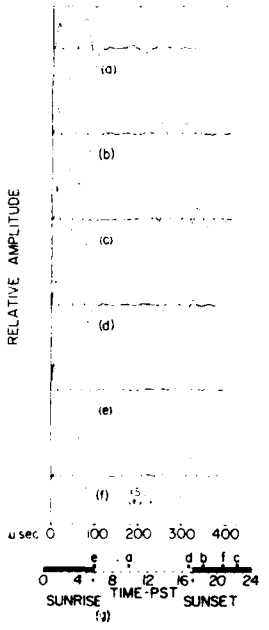



Figure 2. Waveform examples for:
(a) Day, (b) Night,
(c) Extended sky wave,
(d) Inverted sky wave,
(e) sunrise sky wave,
(f) two layer reflections,
(g) Time-of-day indicator.

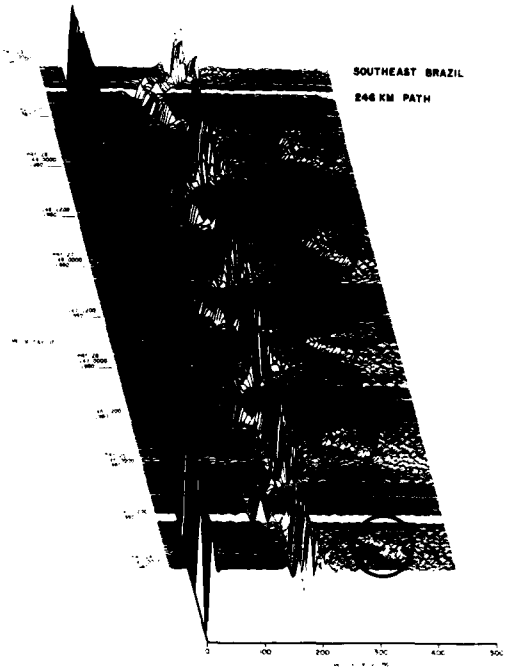



Figure 3. Example of VLF/LF pulse ionosounder data in a three-dimensional format. The circles show moving reflections at high altitudes observed at night; the square indicates weak daytime reflections coming from below the classical D region.

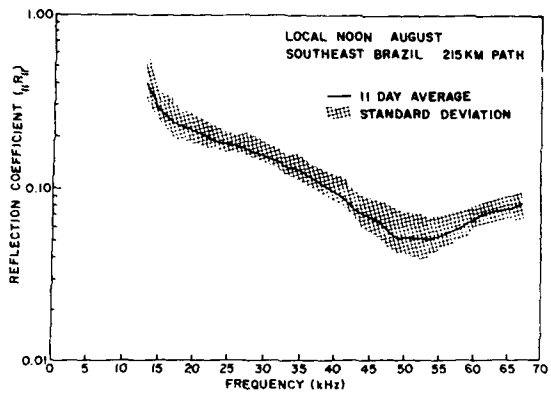



Figure 4. Example of reflection coefficients derived from data acquired at noon over a eleven day period.

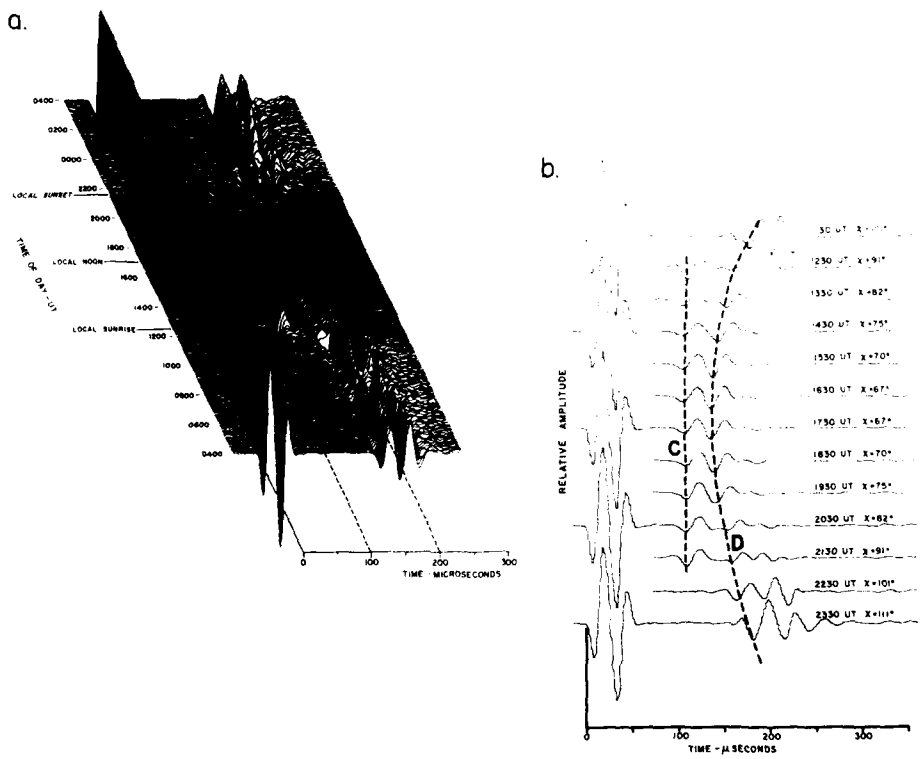



Figure 5. (a) Mid-latitude data obtained on December 28, 1972 over a 263 km path, (b) Examples of waveforms as a function of time-of-day and solar zenith angle.

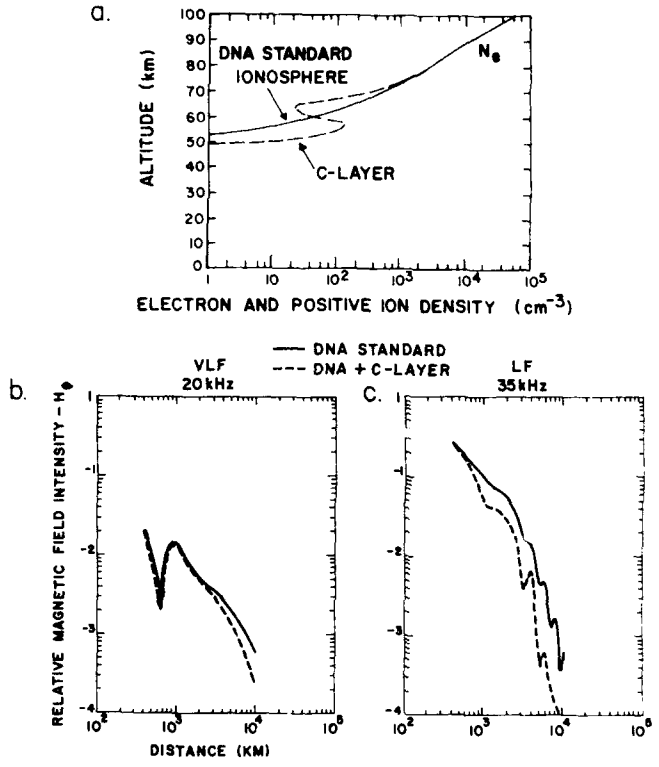



Figure 6. C-layer effects on long wave propagation. (a) DNA daytime ionosphere and C layer model (Bain, 1974) used to calculate relative field strengths versus distance for (b) 20 kHz, and (c) 35 kHz.

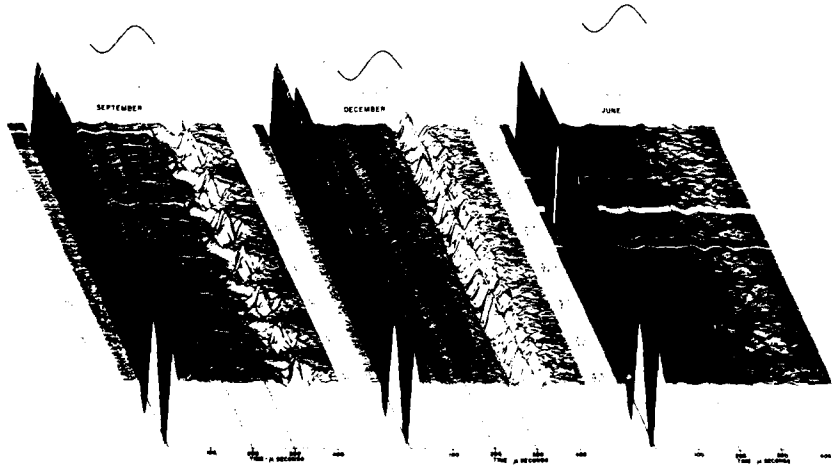



Figure 7. Pulse reflection data typically received under quiet polar ionospheric conditions.

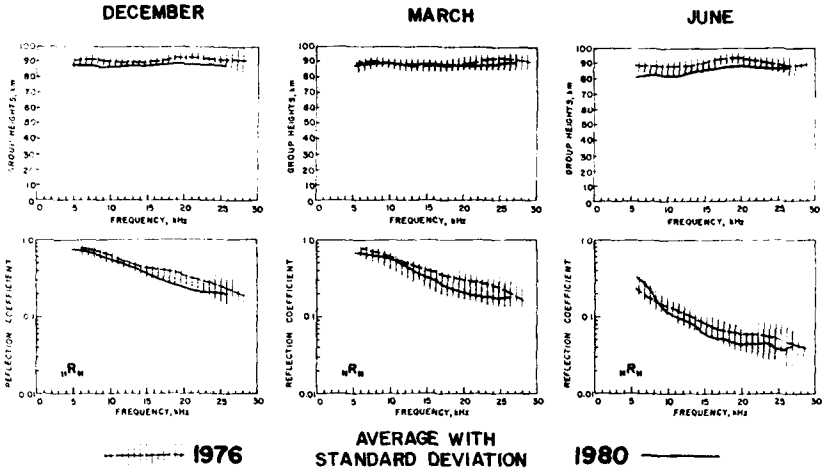



Figure 8. Summary of VLF reflection heights and reflection coefficients derived from pulse reflection data obtained at local midnight under quiet ionospheric conditions in 1976 and 1980.

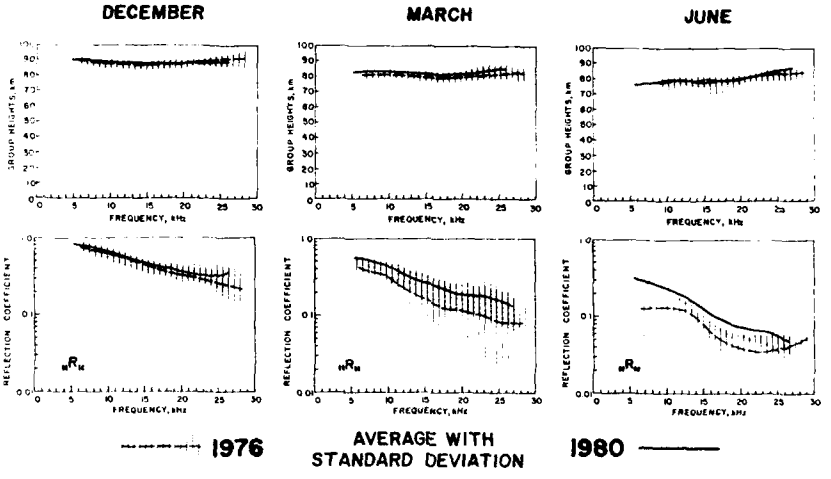



Figure 9. Summary of VLF reflection properties of the quiet polar ionosphere derived from data obtained at local noon in 1976 and 1980.

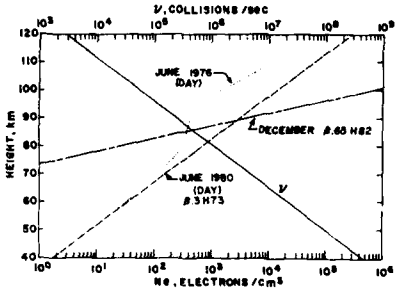



Figure 10. Electron density profiles of the quiet polar ionosphere derived from VLF reflection data.

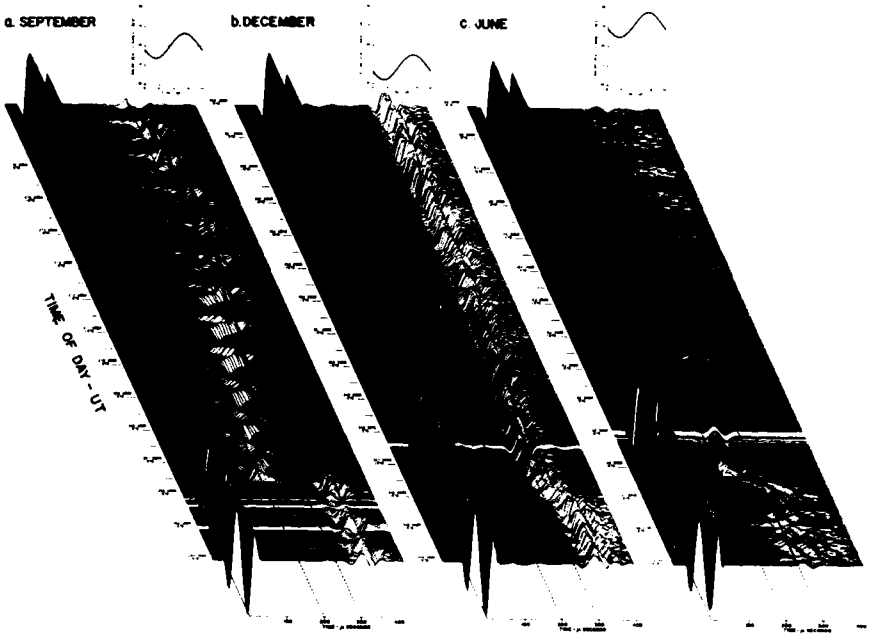



Figure 11. VLF pulse reflection data obtained during energetic particle events, under different solar illumination conditions.

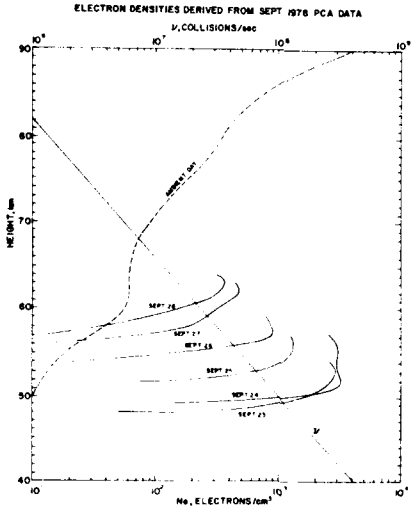



Figure 12. Electron density profiles derived from VLF reflection data obtained at noon during the September 1978 energetic particle event.

## DISCUSSION

**G.Rostoker**

From the AFCRL Handbook, I understand that the layer of ionization produced by cosmic rays is found around altitudes of 10-15 km, with a rapid falloff in charge density with increasing altitude. I find this hard to reconcile with your $C$ layer near 50 km. Can you explain this to me?

**Author's Reply**

I believe you are referring to the fact that the production rate of ion-pairs by galactic cosmic rays peaks in the 10-15 km altitude range. The resultant free electron densities that are produced at such low altitudes, however, are negligible, owing to extremely high electron-neutral collision frequencies. The heavy (negative) ions that are formed by the collision process do not appreciably affect the propagation of VLF/LF waves. On the other hand, data described in this paper provide clear evidence that a sufficient number of free electrons can be maintained at around 60 km altitude, by the combined action of cosmic rays and solar illumination, to produce observable effects on the propagation of low frequency radio waves.

# THE MIDDLE AND HIGH LATITUDE IONOSPHERE AT ∿ 550 km ALTITUDE

Y.K. TULUNAY
İ.T.Ü. Uçak ve Uzay Bilimleri Fakültesi
Maslak, İstanbul, Turkey

SUMMARY

Ambient electron density measurements made by the radio frequency capacitance probe on Ariel 4 satellite have been analysed in order to study the structures present at the satellite altitudes, i.e. between ∿ 470 and ∿ 600 km, at latitudes poleward of ± 50° Λ both during winter and equinox; data obtained on 94 days centered on the solstices, and March 1972 have been used. The general morphology of and the extreme densities, the mid-latitude electron density trough, the polar cap depletions and the electron density enhancements associated with the cusp auroral zone were determined statistically. Analyses of the solstice and the equinox data acquired during quiet magnetic conditions ($K_p \leq 2+$) show that the Northern and Southern Hemisphere ionospheres were significantly different. The global characteristics of the winter solstice and the equinox electron densities were similar although they differ from each other in fine details.

## 1. INTRODUCTION

Tulunay and Grebowsky [1] and Tulunay [2] after analysing electron density data obtained by means of the Ariel 4 Satellite during two magnetically quiet ($K_p \leq 2+$) periods centered on the Winter solstice and the march 1972 equinox at high latitudes in both the Northern and Southern Hemispheres, reported that the Northern and the Southern Hemisphere winter and equinox ionospheres were significantly different. The analysis technique used by Brinton et al. [3] has been adopted in both of these studies. In this paper the results of the above mentioned papers are reported in order to facilitate a simultaneous comparative examination of the earth's ionosphere at about 550 km.

## 2. THE ARIEL 4 SATELLITE AND THE DATA ANALYSIS TECHNIQUE

Ariel 4 was launched in December 1971 into a near-circular orbit having an inclination of 83°; the orbital period was 95 min, with perigee and apogee occurring at approximately 470 and 600 km, respectively. The electron density data were obtained by means of a radio-frequency capacitance probe [4] and all the results used were obtained from tape-recorded data; measurements of the electron density were made at latitude intervals of 1.9° ± 0.5° at low and middle latitudes [5]. The winter periods investigated were from 16 December 1971 to 19 March 1972 for the Northern Hemisphere and from 26 April to 29 July 1972 for the Southern Hemisphere. The equinox period investigated was from 4 February 1972 to 8 May 1972 for both the Northern and Southern Hemispheres. All the periods corresponded to magnetically quiet periods having the 3-h planetary magnetic activity index $K_p \leq 2+$.

The maximum and minimum electron densities at any given location were obtained using a technique similar to that described by [3]. The electron densities observed by Ariel 4 ranged between about $9\times10^{2}$ and $2\times10^{5}$ $cm^{-3}$ and the actual values were divided into the following seven ranges: (1) $10^{2}$-$4\times10^{2}$, (2) $4\times10^{2}$-$8\times10^{2}$, (3) $8\times10^{2}$-$1.6\times10^{4}$, (4) $1.6\times10^{4}$-$3.2\times10^{4}$, (5) $3.2\times10^{4}$-$6.4\times10^{4}$, (6) $6.4\times10^{4}$-$1.3\times10^{5}$ and (7) greater than $1.3\times10^{5}$ $cm^{-3}$ during the solstice period. The electron densities observed by Ariel 4 ranged between about $1\times10^{3}$ and $3\times10^{5}$ $cm^{-3}$ and the actual values were divided into the following six ranges: (1) $10^{3}$-$8\times10^{3}$, (2) $8\times10^{3}$-$1.6\times10^{4}$, (3) $1.6\times10^{4}$-$3.2\times10^{4}$, (4) $3.2\times10^{4}$-$6.4\times10^{4}$, (5) $6.4\times10^{4}$-$1.3\times10^{5}$, (6) greater than $1.3\times10^{5}$ $cm^{-3}$ during the equinox period. The spatial locations of the measurements made within each of the electron density ranges were plotted on "maps" based on magnetic local time and invariant magnetic latitude, i.e. in M.L.T. -Λ space. By overlaying the plots corresponding to all six electron density ranges in order of decreasing number density, it is possible to obtain the "extreme electron density" topographical maps shown in Fig. 1a or Fig. 2a for both hemispheres; these maps show, for any M.L.T. - Λ location, the maximum electron density observed during the 94-day period. Similarly, by overlaying the six plots in order of increasing electron density, it is possible to obtain topographical maps (Fig. 1b or Fig. 2b) which indicate the minimum electron density observed at each M.L.T. -Λ location.

## 3. RESULTS AND CONCLUSIONS

The main purpose of this paper is to present a comprehensive topographical picture of the extreme ambient electron densities observed over the high magnetic latitudes at about 550 km altitude, using a similar approach to that adopted by Brinton et al. [3] for the ion composition to 300 km. Figures 1a and 2a show the topographical maps of the Northern Hemisphere and Southern Hemisphere electron densities indicating the maximum electron densities measured when $K_p \leq 2+$ at all M.L.T. - Λ locations during the relevant 94-day winter solstice and March equinox periods respectively. Figures 1b and 2b are the topographical maps of the Northern Hemisphere and the Southern Hemisphere electron densities indicating the minimum electron densities during quiet magnetic conditions ($K_p \leq 2+$) at all M.L.T. - Λ locations during the relevant 94-day winter solstice and March equinox periods. The following results and conclusions were reached concerning magnetically quiet conditions ($K_p \leq 2+$) during the 1972 winter solstice periods [1].

(i) The highest electron densities were observed in a region that is symmetrical with respect to the 02-14 M.L.T. meridian in the Northern Hemisphere but symmetrical with respect to the 00-12 M.L.T. meridian in the Southern Hemisphere.

(ii) Electron densities were generally smaller in the Southern Hemisphere than in the Northern Hemisphere.

(iii) The mid-latitude trough is the main structure at night in both hemispheres. In the Northern Hemisphere the low densities did not persist into the dayside, but in the Southern Hemisphere the region of reduced electron density extended almost throughout the day and towards higher latitudes.

(iv) A polar cavity is observed in both hemispheres.

(v) In the Southern Hemisphere the highest winter electron densities observed ($> 1.3 \times 10^5$ cm$^{-3}$) occurred only in the dayside cups region whereas in the Northern Hemisphere the maximum density region extended to low latitudes in the daytime.

Data for the 1972 March equinox periods revealed the following results and conclusions during quiet magnetic conditions ($K_p \leq 2+$) [2].

(i) The highest electron densities were observed in a region that is symmetrical with respect to the 02-15 M.L.T. meridian in the Northern Hemisphere but symmetrical with respect to the 00-12 M.L.T. meridian in the Southern Hemisphere.

(ii) Electron densities were generally smaller in the Southern Hemisphere than in the Northern Hemisphere. The extreme values of the electron densities were also observed in the Southern Hemisphere.

(iii) The mid-latitude trough is the main structure at night in both hemispheres.

(iv) A polar cavity is observed in both hemispheres.

(v) In the Southern Hemisphere the highest winter electron densities observed ($> 1.3 \times 10^5$ cm$^{-3}$) occurred only in the dayside cusp region whereas in the Northern Hemisphere the maximum density region extended to low latitudes between 06 and 18 M.L.T.

A comparison of the winter solstice results with the March equinox results revealed that electron densities were smaller in the winter data in both hemispheres in general. However, both sets of data exhibited a similar pattern statistically on the M.L.T. - $\Lambda$ space. The results showed that during quiet magnetic conditions ($K_p \leq 2+$) the Northern and the Southern Hemisphere winter and equinox ionospheres were significantly different ; in particular, the Southern Hemisphere densities were lower than those in the Northern Hemisphere during both 94-day periods.

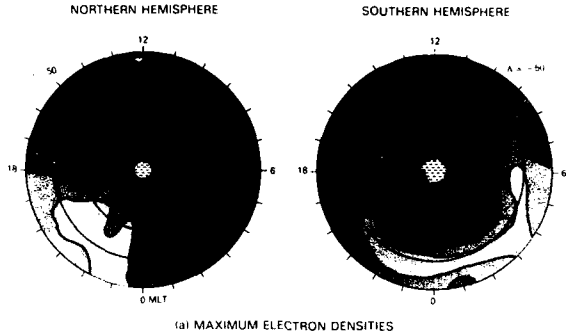



Fig. 1a Topographical "maps" of the Northern Hemisphere (left-hand diagram) and the Southern Hemisphere electron densities indicating the maximum electron densities measured when $K_p \leq 2+$ at all M.L.T.- locations during the relevant 94-day solstice periods.

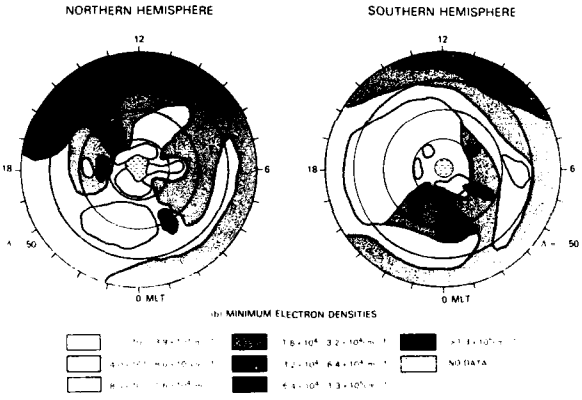



Fig. 1b Topographical "maps" of the Northern Hemisphere (left-hand diagram) and the Southern Hemisphere electron densities indicating the minimum electron densities measured when $K_p \leq 2+$ at all M.L.T.- locations during the relevant solstice periods.

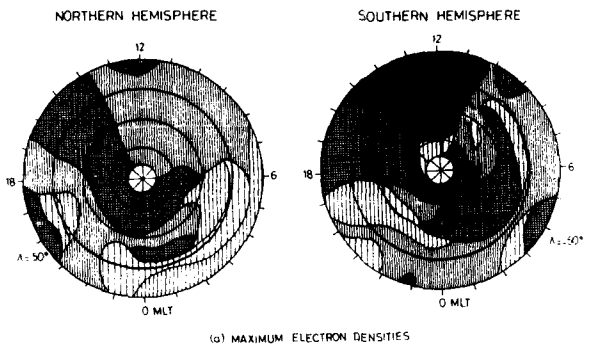



Fig. 2a Topographical "maps" of the Northern Hemisphere (left-hand diagram) and the Southern Hemisphere electron densities indicating the maximum electron densities measured when $K_p \leq 2+$ at all M.L.T.- $\Lambda$ locations during the relevant 94-day equinox periods.

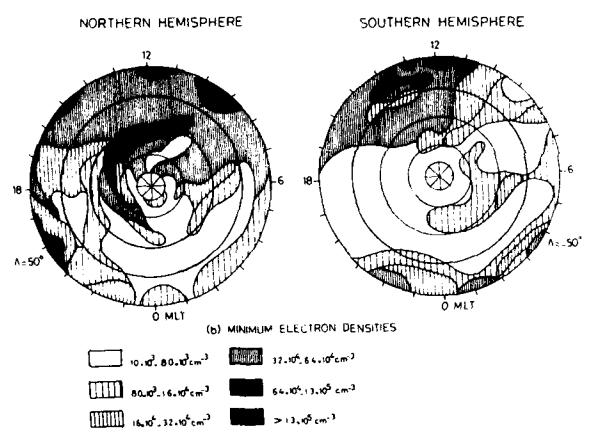



Fig. 2b Topographical "maps" of the Northern Hemisphere (left-hand diagram) and the Southern Hemisphere electron densities indicating the minimum electron densities measured when $K_p \leq 2+$ at all M.L.T.- $\Lambda$ locations during the relevant equinox periods.

## DISCUSSION

**T.Croft**

On August 15, 1972, there was a historically violent solar outburst after which there was virtually no solar wind for three days. Did you observe any ionospheric response to that outburst?

**Author's Reply**

Unfortunately, the experiment went off on that occasion. Therefore, no data were available.

**P.Vila**

(1) Your measurements with 94 days for a diurnal LT cycle need correction if we want to derive $N_m F_2$ distributions. It would be useful to take into account the height of measurements in order to do this.

(2) So if you looked at $N_m F_2$ from ionosonde data you could re-trace the average profile down and correct your set of non-peak data.

(3) Appleton World Data Center should help you!

**Author's Reply**

(1) I evaluated the scale heights corresponding to the height-sampling.

(2) I have no ground data.

**K.Bibl**

Could you define the selection of the minimum and maximum maps?

**Author's Reply**

(i) The arithmetic means of the electron densities (ED) obtained during 2-hr MLT intervals have been studied as a function of invariant magnetic latitudes, $\Lambda$.

(ii) The spatial locations of the measurements made within each of the ED ranges (7 in the case of solstices) were plotted on maps based on MLT-$\Lambda$ space.

(iii) By overlaying the plots corresponding to all seven electron density ranges in order of decreasing number density, the maximum electron density topographical maps for both hemispheres were obtained.

(iv) Similarly, by overlaying the plots corresponding to all seven electron density ranges in order of increasing number density, the minimum electron density topographical maps for both hemispheres were obtained.

# MODELLING IONOSPHERIC DENSITY STRUCTURES

R. W. Schunk and J. J. Sojka

Center for Atmospheric and Space Sciences
Utah State University
Logan, Utah 84322-4405 USA

## SUMMARY

Large-scale density structures are a common feature in the high-latitude ionosphere. They have been observed in the dayside cusp, polar cap, and nocturnal auroral region over a range of altitudes, including the $E$-region, $F$-region and topside ionosphere. The origins, lifetimes and transport characteristics of large-scale density structures were studied with the aid of a three-dimensional, time-dependent ionospheric model. The study considered blob creation due to particle precipitation, the effect that structured electric fields have on the ionosphere, and the lifetimes and transport characteristics of density structures for different seasonal, solar cycle, and interplanetary magnetic field (IMF) conditions. The main conclusions drawn from the study are: (1) The observed precipitation energy fluxes are sufficient for blob creation if the plasma is exposed to the precipitation for 5–10 minutes; (2) Structured electric fields produce structured electron densities, ion temperatures, and ion composition; (3) The lifetime of an $F$-region density structure depends on several factors, including the initial location where it was formed, the magnitude of the perturbation, season, solar cycle and IMF; and (4) Depending on the IMF, horizontal plasma convection can cause an initial structure to break up into multiple structures of various sizes, remain as a single distorted structure, or become stretched into elongated segments.

## 1. INTRODUCTION

During the last decade, significant progress has been made in understanding the formation and evolution of the major morphological features observed in the high latitude ionosphere. This understanding has been reflected in the ability of the ionospheric models to successfully describe such features as the mid-latitude trough, the polar hole, the cross-polar cap tongue of ionization, and the auroral ionization enhancements. The ionospheric models have also been successful in describing the variation of these features with solar cycle, season, and geomagnetic activity (convection pattern) and in describing the universal time (UT) dependence of these features that results from the displacement between the geomagnetic and geographic poles [1–11].

More recently, considerable attention has been directed towards studying the plasma density structures and irregularities that have been observed at high latitudes [12–25]. There are small-scale ($\lesssim$ 1 km), medium-scale ($\sim$ 10 km), and large-scale ($\gtrsim$ 100 km) density structures. These structures have been observed in the $E$-region, $F$-region and topside ionosphere throughout the polar region, including the dayside cusp, polar cap, and auroral zone. The small-scale structures appear to be produced within and on the edges of the larger structures through plasma instabilities and are typically referred to as density irregularities [26, 14, 27]. The medium and large-scale structures have been called 'blobs', 'patches', and 'enhancements'.

Numerous structuring mechanisms have been proposed, including structured precipitation, structured electric fields, structured heat flow, transport from distant sources, structured enhanced decay, as well as time varying parameters of all kinds [25, 12, 21, 18, 14, 24]. Although all of the proposed mechanisms can produce irregularities in a qualitative sense, very little attention has been given to detailed quantitative studies of irregularity formation. However, it is important to understand irregularity production and transport mechanisms because plasma enhancements affect the momentum and energy coupling between the ionosphere and thermosphere. They also increase the height-integrated conductivity and thereby affect the electrical coupling of the ionosphere-magnetosphere circuit. Consequently, we have undertaken a systematic quantitative study of the origins, lifetimes, and transport characteristics of medium and large-scale density structures using a time-dependent three-dimensional ionospheric model. The study considered blob creation due to particle precipitation, the effect that structured electric fields have on the ionosphere, and the lifetimes and transport characteristics of plasma density structures for different seasonal, solar cycle, and interplanetary magnetic field conditions.

## 2. IONOSPHERIC MODEL

The ionospheric model was initially developed as a mid-latitude, multi-ion ($NO^+$, $O_2^+$, $N_2^+$, and $O^+$) model by [28]. The time-dependent ion continuity and momentum equations were solved as a function of altitude for a corotating plasma flux tube including diurnal variations and all relevant $E$ and $F$ region processes. This model was extended to include high latitude effects due to convection electric fields and particle precipitation by [29, 30]. A simplified ion energy equation was also added, which was based on the assumption that local heating and cooling processes dominate (valid below 500 km). Flux tubes of plasma were followed as they moved in response to convection electric fields. A further extension of the model to include the minor ions $N^+$ and $He^+$, an updated photochemical scheme, and the MSIS atmospheric model is described in [31].

The addition of plasma convection and particle precipitation models is described in [3, 4]. More recently, the ionospheric model has been extended by [32] to include ion thermal conduction and diffusion-thermal heat flow, so that the ion temperature is now rigorously calculated at all altitudes between 120–1000 km. The adopted ion energy equation and conductivities are those given by [33]. Also, the electron energy equation has been included recently by [11] and consequently, the electron temperature is now rigorously calculated at all altitudes. The electron energy equation and the heating and cooling rates were taken from [34], and the conductivities were taken from [35]. The incorporation of the *Sterling et al.* [36] equatorial ionospheric model and the various improvements to this model are described in [10].

In this study, the emphasis is on ionospheric structure at high latitudes. With the high latitude model, flux tubes of plasma are followed as they convect through a moving neutral atmosphere. Altitude profiles of the ion and electron temperatures and the $NO^+$, $O_2^+$, $N_2^+$, $N^+$, and $O^+$, and $He^+$ densities are obtained by solving the appropriate continuity, momentum, and energy equations including all of the high latitude processes thought to be important. The equations are typically solved over the altitude range from 120 to 800 km, with boundary conditions specified at the lower and upper ends

of the flux tubes. For the densities, chemical equilibrium is assumed at 120 km and no escape flux is assumed at 800 km. For the temperatures, local thermal coupling is assumed at 120 km and a specified heat flux is assumed at 800 km. For this study, the electron and ion heat fluxes through the upper boundary were taken to be zero.

## 3. PARTICLE PRECIPITATION IN THE DAYSIDE OVAL

The effect of particle precipitation on the polar cap ionosphere was studied by [37]. In particular, a flux tube of plasma was followed as it convected across the polar cap from the dayside to the nightside. For a brief period of time on the dayside, it was subjected to enhanced ionization owing to electron precipitation in the dayside oval and the effect on the electron density profile was calculated using a time-dependent ionospheric model.

Figure 1 shows the four $O^+$ production rate profiles used in the study. The solid line is for a hard auroral precipitation spectrum. The spectrum corresponds to an auroral energy flux of 0.92 ergs $cm^{-2}$ $s^{-1}$ $sr^{-1}$. Soft, subkilovolt, auroral $O^+$ production is represented by the dashed curve labelled 1. This profile corresponds to cusp precipitation and has an associated auroral energy flux of 0.20 ergs $cm^{-2}$ $s^{-1}$ $sr^{-1}$. The average primary energy of the soft and hard auroral spectra are 0.2 and 2 keV, respectively. To simulate the effect of high altitude production, the peak of the "soft" $O^+$ production rate profile 1 was raised by 100 km (curve 2), and then by 300 km (curve 3). The latter case is intended to represent a hypothetical $O^+$ production rate due to extremely high fluxes of very cold precipitating electrons ($\sim$ 50 eV).

Figure 2 shows part of the selected flux tube trajectory in a magnetic latitude-MLT reference frame, from 1200 to 2400 MLT. After midnight this flux tube essentially corotates back to noon. Between steps 16 and 20, the trajectory passes through the region of enhanced auroral ion production. The lower part of Figure 2 shows how the precipitating energy flux varies during the auroral crossing, which takes less than 10 minutes. A Volland convection pattern with enhanced flow in the dusk sector and a cross-tail potential of 62 kV was used to obtain the trajectory shown in Figure 2.

Table 1. Selected Auroral Production Models.

| Trajectory | Description of Precipitation | Peak $O^+$ Altitude, km | Scaling Factor | Energy Flux, erg $cm^{-2}$ $s^{-1}$ $sr^{-1}$ | Plotting Symbol |
|---|---|---|---|---|---|
| 1 | none | - | 0 | - | |
| 2 | hard | 135 | 4 | 3.7 | |
| 3 | soft | 280 | 4 | 0.8 | |
| 4 | raised soft | 380 | 4 | — | |
| 5 | ultra soft | 580 | 4 | - | |
| 6 | ultra soft | 580 | 16 | | |

In Table 1 the various auroral precipitation spectra used for six otherwise identical trajectory runs are shown. For the first trajectory run, no auroral precipitation was included, and consequently, this run can be used as a baseline against which comparisons can be made. Trajectory runs 2 and 3 correspond to the hard and soft (1) production profiles shown in Figure 1, which have been scaled up to give energy fluxes of 3.7 and 0.8 ergs $cm^{-2}$ $s^{-1}$ $sr^{-1}$, respectively. Similarly, trajectory runs 4 and 5 correspond to the soft (2) and (3) production profiles shown in Figure 1 with the energy fluxes scaled by a factor of 4. The final trajectory run 6 is identical to trajectory run 5, except that the energy flux has been scaled by a further factor of four. Also included in Table 1 are the altitudes of the peak $O^+$ production. Both the hard and soft (2 and 3) spectra have production profiles with peaks below the nominal $F$-region peak ($\sim$ 350 km). The results of following the trajectory shown in Figure 2 are contrasted for each of the dayside auroral ionization sources of Table 1 in the following paragraphs.

Figure 3 shows how $N_mF_2$ varies around the trajectory for the six cases. Prior to entering the dayside oval (step 16), the $N_mF_2$ variations for the six cases are the same. In the arc (between steps 16 and 20), the $N_mF_2$ variations are different because the six flux tubes are subjected to different ionization sources. Also, for a short period of time after the flux tubes leave the precipitation region, the $N_mF_2$ variations are different (between steps 20 and 30); this is the time it takes for the flux tubes to adjust to the absence of precipitation. However, when the flux tubes are in darkness (steps 30–80), the $N_mF_2$ variations are the same. Only after the flux tubes enter sunlight (step 80) do the $N_mF_2$ values merge together. This indicates that any dayside enhancement that reaches the terminator before being destroyed will maintain its size relative to the background plasma throughout the night. Likewise, any enhancement created on the nightside will be maintained. Only when the enhancement and the background plasma enter a region of strong uniform production will the enhancement disappear.

Figure 4 shows how $h_mF_2$ varies around the trajectory for the six cases. The important feature to note is that only in a very restricted region are the $h_mF_2$ values different (between steps 16 and 30). In the arc (steps 16–20), the $h_mF_2$ variations are different because the six flux tubes are subjected to ionization sources that peak at different altitudes. The hard precipitation source produces a low $h_mF_2$ ($\sim$ 250 km), while the soft precipitation source that peaks at high altitudes produces a high $h_mF_2$ ($\sim$ 500 km). However, once the flux tubes leave the precipitation region, $h_mF_2$ quickly readjusts to its nominal value determined by diffusion and induced vertical drifts. By about step 25, which is only 25 min after the flux tubes leave the auroral region, the $h_mF_2$ variations are nearly the same. Note that beyond step 30, the $h_mF_2$ values are the same for all six cases even though the $N_mF_2$ values are different. This indicates that outside the source region, $N_mF_2$ and $h_mF_2$ are not necessarily related. Once an enhancement is created, temporal and spatial changes in the convection pattern and/or thermospheric wind will affect $h_mF_2$ even if $N_mF_2$ doesn't change. Consequently, the fact that an enhancement is measured with a high $h_mF_2$ is not necessarily indicative of the source if the enhancement is outside of the source region.

## 4. STRUCTURED ELECTRIC FIELDS

Structured electric fields are a common phenomenon in the polar cap when the IMF is northward [38–41]. These structures are generally superimposed on a large-scale pattern of plasma convection. The spatial scale of the structured electric fields ranges from 10's of kilometers to 1000 km, with the average size being a few hundred kilometers. The electric field variation from the average value ranges from a few mV/m to more than 100 mV/m. The effect that these structured

electric fields have on the polar cap ionosphere was studied by [42] using a three-dimensional time-dependent ionospheric model.

In order to model a convection pattern with electric field structure superimposed, measurements are needed not only along the satellite track, but also in a direction perpendicular to the track so that the entire spatial extent of the structure can be determined. Unfortunately, the required data set was not available and the model study [42] had to be based on a semi-empirical description of electric field structure. The structure was assumed to be associated with a discrete current system into and out of the polar ionosphere; the currents sheets for each structure were assumed to be sun-aligned and had a finite length. For lack of measurements, the potential disturbance associated with the structured current sheets was assumed to be a sun-aligned Volland two-cell pattern [43], as shown in Figure 5. Each electric field structure was described by one such 2-cell pattern, but the spatial size of the structure and potential drop across it varied from structure to structure. In the model study, a total of 40 structures were randomly distributed in the polar cap, and a Gaussian random number generator was used to select the parameters that are needed to define a single structure. In using the random number generator, the parameters were assumed to have a Gaussian distribution about a 'specified' mean value. Figure 6 shows the structure distribution for the case where the electric potential drops across the central regions of the structures have a mean value of 12 kV and a Gaussian half-width of 4 kV.

The structured potential distribution shown in Figure 6 was added to a standard, large-scale Volland potential distribution, which was used to simulate the background convection pattern. This background pattern was taken to be a symmetric two-cell pattern with a cross polar cap potential drop of 60 kV. In the model study, 8 representative convection trajectories were selected and flux tubes of plasma were followed as they convected along these trajectories across the polar cap in an antisunward direction. The calculations were done both with and without the electric field structure. The top panel of Figure 7 shows the trajectories for the unperturbed or background convection pattern, while the bottom panel shows the resulting trajectories when the structured electric fields are superimposed on the background pattern. The electric fields that would be measured by a satellite crossing these convection patterns along the dusk-dawn magnetic meridian are shown in Figure 8. The top panel shows the electric field signature of a classical two-cell convection pattern, while the bottom panel shows the kind of structure that is frequently observed when the IMF is northward.

In the calculations, the trajectories were all started from daytime steady state conditions at the locations marked by the dot in Figure 7, and the response of the plasma was calculated as the flux tubes followed their convection paths. Figure 9 shows how $N_mF_2$ varies along each trajectory for both the unperturbed (top panel) and structured (bottom panel) convection patterns. Note that the densities are plotted versus the trajectory $x$-axis (see Figure 5). For the unperturbed pattern, the plasma following the 8 trajectories convects across the polar cap and eventually into darkness, during which time $N_mF_2$ decays. However, the important feature to note is that the $N_mF_2$ variations for all 8 trajectories are very similar. In sharp contrast, the $N_mF_2$ variations along the 8 trajectories that contain structured electric fields are very different. Upon reaching the nightside auroral zone, the $N_mF_2$ values for the different flux tubes can differ by as much as an order of magnitude. The differences are a result of different transpolar crossing times and different 'localized' electric field strengths. Localized enhancements in the electric field produce ion temperature enhancements and increased $O^+ \rightarrow NO^+$ conversion rates [29], and this, in turn, yields lower $N_mF_2$ values.

## 5. LIFETIME AND TRANSPORT OF LARGE DENSITY STRUCTURES

The lifetime and transport characteristics of 'large' ionospheric density structures (factors of 10 to 100) were studied by [44]. Both density depletions and enhancements were considered. In this study, a density structure was created at a specific location in the high-latitude $F$-region and the subsequent evolution was followed for different seasonal and solar cycle conditions as well as for different orientations of the IMF, i.e., different convection patterns. However, only two 'structure' locations were considered, one on the dayside and one in the dusk sector.

For the simulations involving the dayside structure location, a standard two-cell convection pattern was adopted. The pattern, which is shown in the top panel of Figure 10, corresponds to a symmetric, two-cell configuration of the *Volland* [43] type with corotation added. This pattern is representative of moderate geomagnetic activity with $Kp = 3$ and a total cross polar cap potential of 62 kV. In the polar cap, the electric field is about 17 mV $m^{-1}$ and the corresponding antisunward convection speed is about 300 m $s^{-1}$. The solid line is an adopted 'test' trajectory, which was used to follow convecting plasma flux tubes in different reference frames. In the middle panel, this test trajectory is followed for $3\frac{1}{2}$ traversals in a geographic inertial frame starting on the dayside between 1000 and 1100 local time (shown by the dot in the upper panel). Three complete traversals of the test trajectory take about 27 hours. Although the $3\frac{1}{2}$ trajectory loops overlap in the magnetic frame, they do not overlap in the geographic inertial frame because the convection pattern is fixed in a magnetic frame and the magnetic pole rotates about the geographic pole. If a plasma blob were following this trajectory, it may or may not survive depending on how long it is exposed to ion production due to either auroral particle precipitation or solar EUV radiation (the terminators for summer and winter solstice are denoted by dashed lines). Therefore, a satellite crossing the polar region may or may not see the plasma blob on successive passes. Finally, the bottom panel of Figure 10 shows the track of the plasma flux tube in a corotating geographic frame, which is the frame of reference in which ground-based measurements are made. In this frame, a blob of plasma drifting into view may appear to be coming from any direction, depending on the location of the observing site.

The test trajectory shown in Figure 10 was followed for $3\frac{1}{2}$ traversals starting on the dayside at the location marked by the dot (285° E longitude). Different seasonal and solar cycle cases were considered as well as different initial density perturbations. A start time of 15:30 UT was selected so that the starting location was sunlit in both summer and winter. As a consequence, the background conditions at the starting location could be obtained from daytime steady state solutions. Figure 11 shows the variation of $N_mF_2$ along the test trajectory for three different density enhancements and for both winter and summer conditions at solar minimum. In each panel, the solid curve corresponds to the natural plasma, while for the dotted and dashed curves the 'initial' normal $O^+$ density profile was multiplied by factors of 10 and 100, respectively. The variations of the solar zenith angle and the precipitating auroral electron energy flux along the test trajectory are shown at the top and bottom of each panel, respectively.

In winter, the flux tube is in darkness most of the time, and therefore, auroral precipitation is the only source of plasma along most of the test trajectory. For the natural case, $N_mF_2$ at the starting location is about $5 \times 10^5$ $cm^{-3}$, but as the flux tube convects across the polar cap in darkness, $N_mF_2$ decreases until it reaches the nocturnal auroral oval. Because the test trajectory has a fairly long residence time in the oval, there are distinct $N_mF_2$ enhancements during each of the three traversals of the oval. When the initial $O^+$ density profile is multiplied by factors of 10 and 100 to create density structures

(blobs), the subsequent plasma decay occurs at the same rate as the background plasma (solid curve) owing to the absence of sunlight. However, this is not true when the blob flux tubes enter the nocturnal auroral oval. Even though the blob densities are much greater than the background density when the flux tubes first encounter the nocturnal oval, auroral precipitation is sufficient to slow the decay of the blobs. Therefore, in winter at solar minimum, the lifetime of a blob with a factor of 10 enhancement relative to the background plasma is about 11 hours, while for a factor of 100 enhancement the lifetime is about 19 hours.

In summer, the temporal evolution of the density structures is similar to that in winter, but there are some quantitative differences. First, for the natural case (solid curve), $N_mF_2$ at the dayside starting location is greater in winter than in summer owing to the well-known winter anomaly. The difference between summer and winter $N_mF_2$'s is smaller in the auroral oval because the adopted precipitation pattern is the same in summer and winter. With regard to the density structures, the important thing to note is that the lifetimes are much shorter in summer than in winter. The blob with a factor of 10 enhancement basically disappears in about 4 hours, while for the factor of 100 enhancement, the decay time is about 11 hours. The faster blob decay rate in summer is a result of both the change in the atmospheric $O/N_2$ density ratio and the greater importance of solar EUV production, which helps to maintain background plasma densities.

In the simulations discussed above, the main emphasis was on the lifetimes of initial 'dayside structures' with a range of enhanced densities. However, for the simulations involving the 'dusk structure' location, the emphasis was on following the breakup of a given structure as it drifted in response to convection electric fields for different IMF conditions. In this latter case, the adopted structure was assumed to have a very large density enhancement of a factor of 100 relative to background $O^+$ densities and its 'initial' location was taken to be in the dusk-midnight sector at 2000 MLT. Six plasma flux tubes along the 2000 MLT meridian were selected to represent the initial width of the structure, with dipole latitude locations of 73.1, 74.3, 74.9, 75.5, 76.1, and 76.7°. These initial flux tube locations are shown in Figure 12 in a magnetic latitude-MLT reference frame. Also shown in Figure 12 is the subsequent spatial evolution of these plasma flux tubes for four different IMF configurations. The different cases were obtained from the *Sojka et al.* [46] convection model with corotation added. The top dials show $B_y$ positive and negative cases for southward IMF, while the bottom dials correspond to the same $B_y$ cases for northward IMF. For southward IMF, two-cell convection patterns exist, while for northward IMF multicell convection exists. For $B_y$ positive, enhanced plasma convection occurs in the dawn sector, while for $B_y$ negative the enhanced plasma flow occurs in the dusk sector. It is apparent from Figure 12 that the IMF conditions have an important effect on the subsequent evolution of a density structure. For the case ($B_z = -18\gamma$, $B_y = 24\gamma$), the six trajectories stay together as they drift in response to convection electric fields. However, for the three other IMF configurations, the group of trajectories separate as they convect, i.e., the plasma structure breaks up into separate smaller structures. For the case ($B_z = -18\gamma$, $B_y = -24\gamma$), one segment of the initial structure stays in the dusk convection cell, while another segment crosses the polar cap and then essentially corotates throughout the morning and afternoon sectors before returning to the polar cap. For northward IMF, the breakup of the structure can be even more complicated because of the multicell nature of convection. For the case ($B_z = 18\gamma$, $B_y = 24\gamma$), one segment of the structure stays in the dusk convection cell, while another segment appears to circulate around the polar cap. For the last case ($B_z = 18\gamma$, $B_y = -24\gamma$), the bulk of the structure drifts around the midnight-dawn-noon sector, and as it does, it undergoes a significant spatial spreading and stretching (tick marks indicate hourly intervals).

Figure 13 shows the trajectories of the above plasma flux tubes in a corotating geographic reference frame. To an observer fixed to the ground, the situation could appear to be very complicated, depending on the IMF. Since different trajectories in the 'initial' structure cross each other at different, but later, times a ground observer might think that the density enhancements observed at different times are unrelated. A further complication occurs because the later crossings of the ground site by a density enhancement could be from different directions. It is apparent from Figure 13 that multiple observing sites and extended observations are needed to unambiguously determine structure dynamics, particularly for northward IMF.

Not only can an initial density structure break up into smaller segments, but the subsequent evolution of the individual segments can be very different. This is shown in Figure 14, where $N_mF_2$ is plotted versus time along five of the trajectories shown in Figure 12 for each of the four IMF cases and for winter and solar maximum conditions. The shading at the bottom of each plot indicates the times the convecting flux tubes are subjected to auroral precipitation. Dotted segments of the curves indicate the flux tubes are in darkness, while solid segments indicate sunlit conditions. Considering first the ($B_z = -18\gamma$, $B_y = 24\gamma$) case (left column), it is readily apparent that the temporal evolution of $N_mF_2$ is very similar for the five trajectories because in this case the structure flux tubes stay together (see Figure 12). However, for the other southward IMF case ($B_z = -18\gamma$, $B_y = -24\gamma$), only two of the five structure flux tubes stay sufficiently close to each other to have similar temporal histories (top two plots). For the other three flux tubes, the $N_mF_2$ variations are different because of different convection speeds along the trajectories, different exposures to auroral precipitation (see bottom of plots), and for one flux tube exposure to sunlight (solid curve segment). Note that the general conclusions reached above for southward IMF are also true for northward IMF (two columns on right). Finally, note that for a given IMF configuration and at a given time, the $N_mF_2$ values can be different by more than an order of magnitude even though the five flux tubes started out under the same conditions.

## 6. CONCLUSIONS

Significant progress has been made during the last decade in modelling the major morphological features of the high latitude ionosphere, including the polar cap 'tongue of ionization', the 'main trough', the 'polar hole', auroral ionization peaks, and ion and electron temperature hot spots. The modelling has also been able to describe the universal time variation of the high latitude ionosphere that is associated with the displacement between the geomagnetic and geographic poles. The more recent modelling efforts have focussed on ionospheric structure, with the emphasis on elucidating structuring mechanisms, lifetimes, and transport characteristics. The main conclusions drawn from our studies are as follows: (1) Plasma blobs or enhancements can be created by the observed precipitation energy fluxes if the plasma is exposed to the particle precipitation for 5-10 minutes; (2) Within 10-20 minutes after a density structure is created by precipitation, the $F$-region recovers its shape, and hence, the source characteristics cannot be determined from the profile shape after this time period; (3) Structured electric fields produce structured electron densities, ion temperatures, and ion composition; (4) The lifetime of an $F$-region density structure depends on the initial location where it was formed, the magnitude of the perturbation, season, solar cycle, and the IMF; and (5) Depending on the IMF (i.e., convection pattern), horizontal plasma convection can cause an initial density structure to break up into multiple structures of various sizes, remain as a single distorted structure, or become stretched into elongated segments.

## 7. ACKNOWLEDGEMENTS

This research was supported by NASA grant NAGW 77 and AFOSR contract F49620-86-C-0109 to Utah State University.

## 8. REFERENCES

[1] Knudsen, W. C., P M. Banks, J. D. Winningham, and D. M. Klumpar, Numerical model of the converting $F_2$ ionosphere at high latitudes, *J. Geophys. Res.*, *82*, 1977, 4784–4792.

[2] Watkins, B. J., A numerical computer investigation of the polar $F$-region ionosphere, *Planet. Space Sci.*, *26*, 1978, 599.

[3] 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*, 1981*a*, 609–621.

[4] Sojka, J. J., W. J. Raitt, and R. W. Schunk, Theoretical predictions for ion composition in the high-latitude winter $F$-region for solar minimum and low magnetic activity, *J. Geophys. Res.*, *86*, 1981*b*, 2206–2216.

[5] Sojka, J. J., W. J. Raitt and R. W. Schunk, Plasma density features associated with strong convection in the winter high-latitude $F$-region, *J.Geophys. Res.*, *86*, 1981*c*, 6908–6916.

[6] Sojka, J. J., R. W. Schunk and W. J. Raitt, Seasonal variations of the high-latitude $F$-region for strong convection, *J. Geophys. Res.*, *87*, 1982, 187–198.

[7] Schunk, R. W., and J. J. Sojka, Ionospheric hot spot at high latitudes, *Geophys. Res. Lett.*, *9*, 1982*b*, 1045–1048.

[8] Sojka, J. J., and R. W. Schunk, A theoretical study of the high-latitude $F$ region's response to magnetospheric storm inputs, *J. Geophys. Res.*, *88*, 1983, 2112–2122.

[9] Sojka, J. J., and R. W. Schunk, A theoretical $F$-region study of ion compositional and temperature variations in response to magnetospheric storm inputs, *J. Geophys. Res.*, *89*, 1984, 2348–2358.

[10] Sojka, J. J., and R. W. Schunk, A theoretical study of the global $F$-region for June solstice, solar maximum, and low magnetic activity, *J. Geophys. Res.*, *90*, 1985, 5285–5298.

[11] Schunk, R. W., J. J. Sojka, and M. D. Bowline, Theoretical study of the electron temperature in the high latitude ionosphere for solar maximum and winter conditions, *J. Geophys. Res.*, *91*, 1986, 12,041–12,054.

[12] 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*, 1980, 789–792.

[13] 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*, 1982, 907.

[14] Kelley, M. C., J. F. Vickrey, C. W. Carlson, and R. Torbert, On the origin and spatial extent of high-latitude $F$ region irregularities, *J. Geophys. Res.*, *87*, 1982, 4469–4475.

[15] 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*, 1983, 995.

[16] 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*, 1985, 325.

[17] Basu, S., E. MacKenzie, S. Basu, H. C. Carlson, D. A. Hardy, F. J. Rich, and R. C. Livingston, Coordinated measurements of low energy electron precipitation and scintillations/TEC in the auroral oval, *Radio Sci.*, *18*, 1983, 1151.

[18] Kelly, J. D., and J. F. Vickrey, $F$-region ionospheric structure associated with antisunward flow near the dayside polar cusp, *Geophys. Res. Lett.*, *11*, 1984, 907–910.

[19] Weber, E. J., and J. Buchau, Observation of plasma structure and transport at high latitudes, in *The Polar Cusp*, edited by J. A. Holtet and A. Egeland, Hingham, Mass., D. Reidel, 1985, p. 279.

[20] Foster, J. C., and J. R. Doupnik, Plasma convection in the vicinity of the dayside cleft, *J. Geophys. Res.*, *89*, 1984, 9107–9113.

[21] 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*, 1984, 1683.

[22] Weber, E. J., R. T. Tsunoda, J. Buchau, R. E. Sheehan, D. J. Strickland, W. Whiting, and J. G. Moore, Coordinated measurements of auroral zone enhancements, *J. Geophys. Res.*, *90*, 1985, 6497.

[23] Weber, E. J., J. A. Klobucher, 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*, 1986, 12,121–12,129.

[24] 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 nighttime $F$ region densities at high latitudes, *J. Geophys. Res.*, *90*, 1985, 4319–4332.

[25] Robinson, R. M., R. T. Tsunoda, J. F. Vickrey, and L. Guerin, Sources of $F$ region ionization enhancements in the nighttime auroral zone, *J. Geophys. Res.*, *90*, 1985, 7533.

[26] Fejer, B. G., and M. C. Kelley, Ionospheric irregularities, *Rev. Geophys.*, *18*, 1980, 401.

[27] Lee, M. C., Ohmic dissipation of Pederson current as the cause of high-latitude $F$ region ionospheric irregularities, *J. Geophys. Res.*, *89*, 1984, 2289.

[28] Schunk, R. W., and J. C. G. Walker, Theoretical ion densities in the lower ionosphere, *Planet. Space Sci.*, *21*, 1973, 1875–1896.

[29] Schunk, R. W., W. J. Raitt, and P. M. Banks, Effect of electric fields on the daytime high-latitude $E$-and $F$-regions, *J. Geophys. Res.*, *80*, 1975, 3121–3130.

[30] Schunk, R. W., P. M. Banks, and W. J. Raitt, Effects of electric fields and other processes upon the nighttime high latitude $F$-layer, *J. Geophys. Res.*, *81*, 1976, 3271–3282.

[31] Schunk, R. W., and W. J. Raitt, Atomic nitrogen and oxygen ions in the daytime high-latitude $F$-region, *J. Geophys. Res.*, *85*, 1980, 1255–1272.

[32] Schunk, R. W., and J. J. Sojka, Ion temperature variations in the daytime high-latitude $F$-region, *J. Geophys, Res.*, *87*, 1982*a*, 5169–5183.

[33] Conrad, J. R., and R. W. Schunk, Diffusion and heat flow equations with allowance for large temperature differences between interacting species, *J. Geophys. Res.*, *84*, 1979, 811–822.

[34] Schunk, R. W., and A. F. Nagy, Electron temperatures in the $F$-region of the ionosphere: Theory and observations, *Rev. Geophys. Space Phys., 16*, 1978, 355–399.

[35] Schunk, R. W., and J. C. G. Walker, Transport properties of the ionospheric electron gas, *Planet. Space Sci., 18*, 1970, 1525–1550.

[36] Sterling, D. L., W. B. Hanson, R. J. Moffett, and R. G. Baxter, Influence of electromagnetic drifts and neutral air winds on some features of the $F_2$ region, *Radio Sci., 4*, 1969, 1005–1023.

[37] 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*, 1986, 3215–3253.

[38] Heppner, J. P., Empirical models of high-latitude electric fields, *J. Geophys. Res., 82*, 1977, 1115–1125.

[39] Burke, W. J., M. C. Kelley, R. C. Sagalyn, M. Smiddy, and S. T. Lai, Polar cap electric field structures with a northward interplanetary field, *Geophys. Res. Lett., 6*, 1979, 21–24.

[40] Frank, L. A., J. D. Craven, D. A. Gurnett, S. D. Shawhan, D. R. Weimer, J. L. Burch, J. D. Winningham, C. R. Chappell, J. H. Waite, R. A. Heelis, N. C. Maynard, M. Sugiura, W. K. Peterson, and E. G. Shelley, The theta aurora, *J. Geophys. Res., 91*, 1986, 3177–3224.

[41] Heelis, R. A., P. H. Reiff, J. D. Winningham, and W. B. Hanson, Ionospheric convection signatures observed by DE 2 during northward interplanetary magnetic field, *J. Geophys. Res., 91*, 1986, 5817–5830.

[42] Sojka, J. J., and R. W. Schunk, A model study of how electric field structures affect the polar cap $F$-region, *J. Geophys. Res.*, in press, 1988.

[43] Volland, H., A model of the magnetospheric electric convection field, *J. Geophys. Res., 83*, 1978, 2695–2699.

[44] Schunk, R. W., and J. J. Sojka, A theoretical study of the lifetime and transport of large ionospheric density structures, *J. Geophys. Res., 92*, 1987, 12,343–12,351.

[45] Spiro, R. W., P. H. Reiff and L. J. Maher, Precipitating electron energy flux and auroral zone conductances, An empirical model, *J. Geophys. Res., 87*, 1982, 8215–8227.

[46] Sojka, J. J., C. E. Rasmussen and R. W. Schunk, An IMF dependent model of the ionospheric convection electric field, *J. Geophys. Res., 91*, 1986, 11,281–11,290.

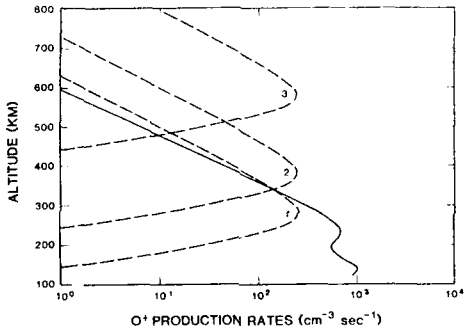



Figure 1. Theoretical $O^+$ production rates as a function of altitude. The solid curve corresponds to the $O^+$ production rate for a typical hard auroral spectrum, while the dashed curve (1) corresponds to a soft cusp spectrum. The dashed curves (2) and (3) are the same as curve (1), except that the profile has been raised to higher altitudes. From [37].

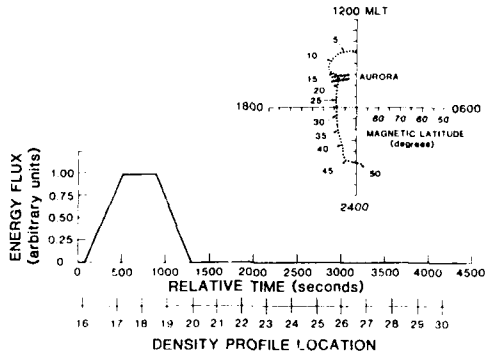



Figure 2. Test trajectory for plasma transport across the polar cap in a magnetic latitude-MLT reference frame (top). Density profile locations have been numbered along the trajectory, and the cross-hatch region corresponds to an auroral ionization source. The bottom part of the figure shows the relation between elapsed time, density profile location, and auroral energy flux. From [37].

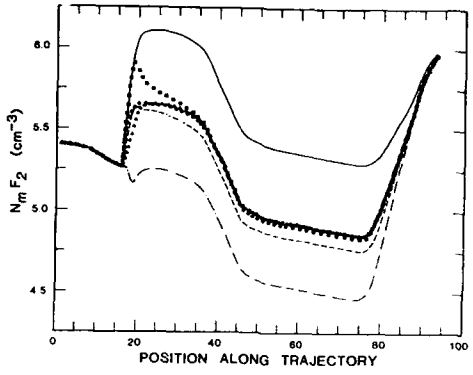



Figure 3. $N_m F_2$ variation along the test trajectory for the different auroral production models defined in Table 1. From [37].

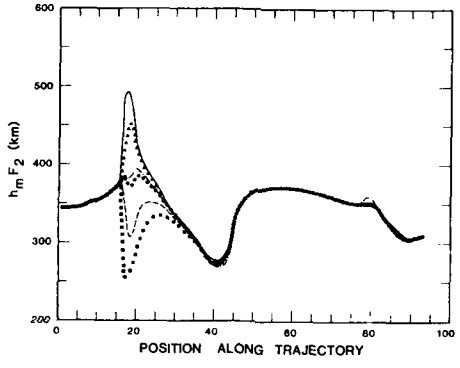



Figure 4. $h_m F_2$ variation along the test trajectory for the different auroral production models defined in Table 1. From [37].

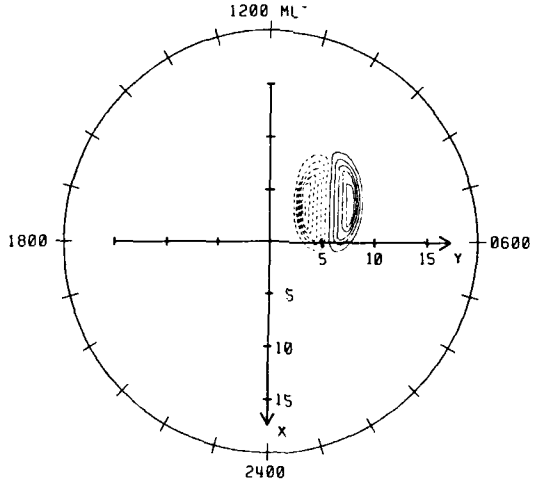



Figure 5. Model of an electric field structure. The electric potential contours drawn at 1 kV intervals are shown in the magnetic $x$-$y$ coordinate system. Negative potentials are shown as dashed contours. The potential varies from $-6$ to 6 kV for this electric field structure. The coordinate system is located at the magnetic pole with the positive $x$-axis lying along the magnetic midnight meridian, while the positive $y$-axis lies along the dawn magnetic meridian. Units along each axis correspond to degrees of magnetic colatitude. From [38].

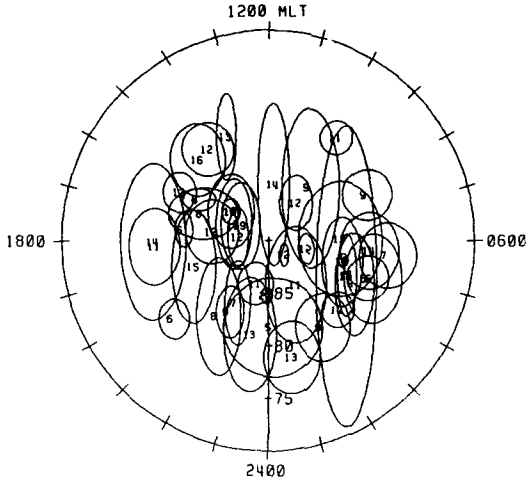



Figure 6. Distribution of the adopted 45 polar cap electric field structures in the MLT-magnetic latitude frame. Each structure is represented by its uniform central region and labeled in the center with its dawn-dusk potential drop in kV. From [38].

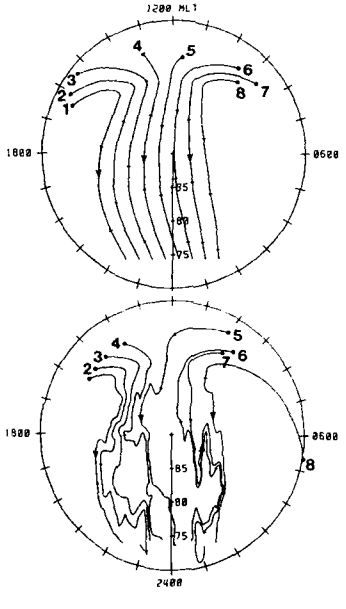



Figure 7. Plasma trajectories for a 2-cell unperturbed convection pattern (top panel) and for a structured convection pattern (bottom panel) in the MLT-magnetic latitude reference frame. Circles denote the start location for each trajectory, while tick marks indicate one-hour time intervals. From [38].

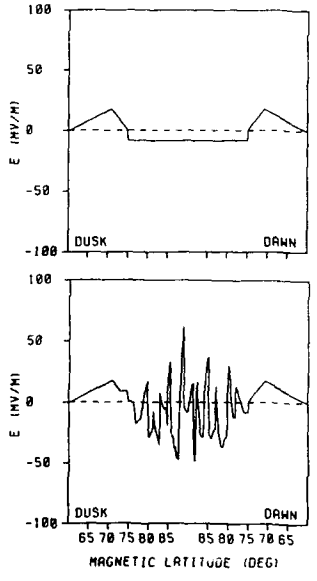



Figure 8. Electric field variation along the dusk-dawn magnetic meridian for the unperturbed 2-cell convection pattern (top panel) and for the structured convection pattern (lower panel). The electric field is given in mV/m. From [38].

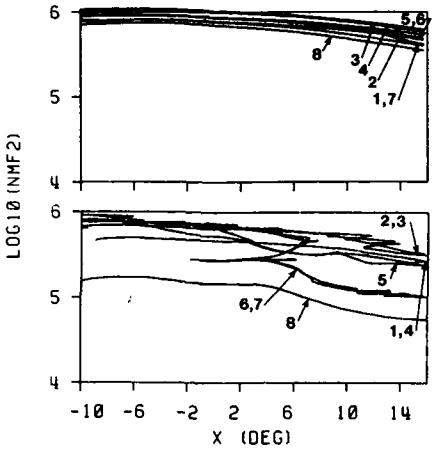



Figure 9. $N_mF_2$ densities for each of the trajectories shown in Figure 7 plotted versus the flux tubes' $x$ location. The top panel is for the unperturbed 2-cell convection pattern and the bottom panel is for the structured convection pattern. $N_mF_2$ is plotted as $\log_{10} N_mF_2$ in $cm^{-3}$. From [38].

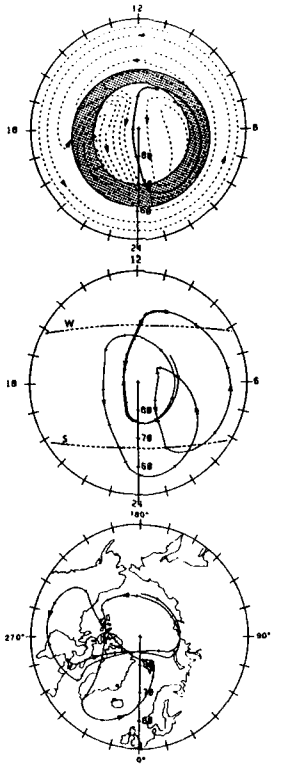


Figure 10. Plasma drift trajectories in different coordinate systems for a *Volland* [43] convection pattern with corotation added. The convection pattern is a symmetric two-cell pattern with a cross polar cap potential of 62 kV. The top panel shows the drift trajectories in a magnetic latitude-MLT reference frame. Also shown in this panel are a *Spiro et al.* [45] auroral oval for $Kp = 3$ and a selected 'test' trajectory (solid curve). The middle panel shows the path of a plasma flux tube in the geographic inertial frame for $3\frac{1}{2}$ traversals of the test trajectory. Also shown in the middle panel are the terminator locations for summer and winter solstice. The bottom panel shows the flux tube path in a corotating geographic reference frame. Tick marks along the flux tube path show the elapsed time at hourly intervals. From [44].

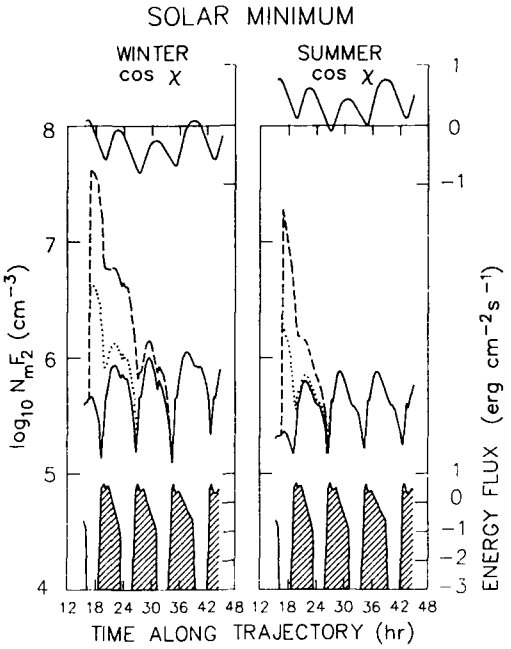



Figure 11. $N_mF_2$ variation along the test trajectory for three plasma flux tubes in both summer and winter at solar minimum. The flux tubes start on the dayside at the location marked by the dot in Figure 10. The solid curve corresponds to the background plasma, while for the dotted and dashed curves the initial $O^+$ density profile was multiplied by factors of 10 and 100, respectively. The solar zenith angle variation along the trajectory is shown at the top, and the log of the auroral electron energy flux along the trajectory is shown at the bottom. From [44].

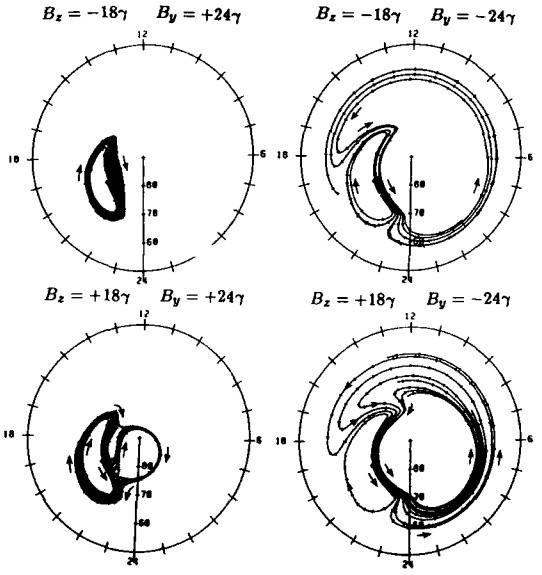



Figure 12. Plasma convection trajectories in a magnetic latitude-MLT reference frame for different IMF conditions. The trajectories were obtained from the *Sojka et al.* [46] convection patterns with corotation added. The top dials show $B_y$ positive and negative cases for southward IMF, while the bottom dials correspond to the same $B_y$ cases for northward IMF. For each of the four IMF cases, six trajectories were followed for 24 hours starting at 2000 MLT and magnetic latitudes of 73.1, 74.3, 74.9, 75.5, 76.1, and 76.7°. The tick marks along the trajectories represent hourly intervals. From [44].

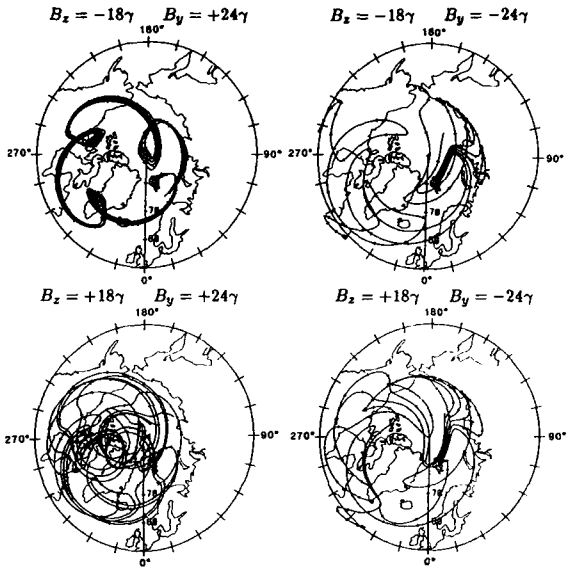



Figure 13. Plasma convection trajectories in the corotating geographic frame for the different IMF cases shown in Figure 12. From [44].

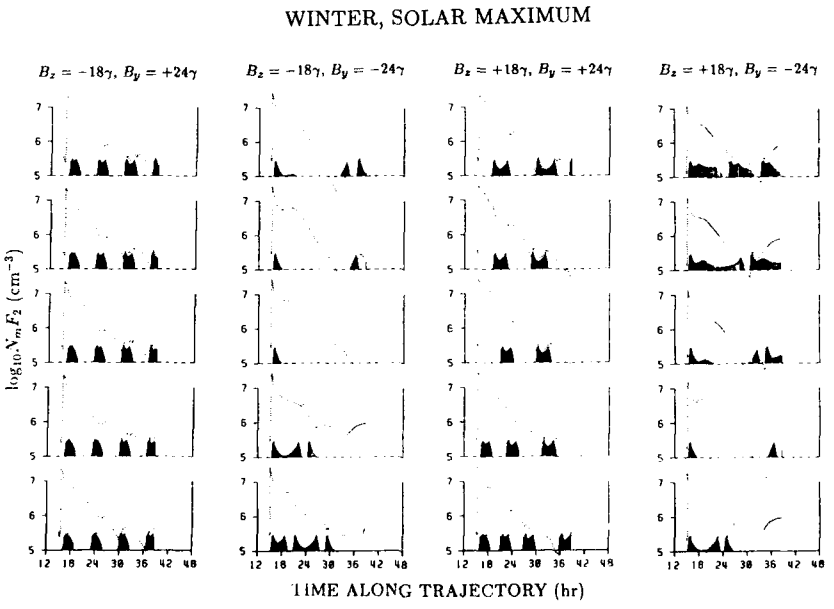



Figure 14. $N_m F_2$ variation along five convection trajectories for four IMF cases in winter at solar maximum. The cases and trajectories are those shown in Figures 12 and 13. The shading at the bottom of each plot indicates the times the flux tubes are subjected to auroral precipitation. Dotted segments of the curves indicate the times the flux tubes are in darkness, while solid segments correspond to sunlit conditions. From [44].

## DISCUSSION

**P.Vila**

About the polar $F$ lacunae: They occur exclusively in summer daytime when the lifetime of the blob-patch fronts is found shortest in your 3D-time simulation.

The existence of simultaneous instabilities at $E$ and $F$ region heights, checked on some meridian scans at Sondrestrom on the front borders, seems to be built in with the electric field and current structure, needing closure by Pedersen currents exclusively on the dayside. Is the imaging $J_{\parallel}$ process of Heelis, Vickrey and Walker [1985] not responsible for blob decay?

**Author's Reply**

In our ionospheric simulations we can only model the dynamics of large-scale patches or blobs. Our model does not allow plasma instabilities to form on the blob edges. Also, our model does not take account of $E$-$F$ imaging processes coupled through field-aligned currents. The $E$-$F$ region image process can speed the decay of small density irregularities through enhanced cross-field diffusion, but it probably is not significant for large-scale patches and blobs.

**B.Reinisch**

For $B_z < 0$, you modelled the convection as a superposition of many small-scale cells, producing a highly structured 2-cell pattern. What is the physical mechanism for these small circulation cells, and what evidence exists?

**Author's Reply**

Each small cell is associated with two small sheet currents into and out of the ionosphere so that the overall structure results from many small sheet currents. Evidence for such a configuration is provided by the Dynamics Explorer data for $B_z$ northward and for conditions when the convection pattern appears to be turbulent.

# OBSERVATIONS OF IONOSPHERE / MAGNETOSPHERE INTERACTIONS FROM THE DYNAMICS EXPLORER SATELLITES

R. A. Hoffman
Laboratory for Extraterrestrial Physics
Goddard Space Flight Center
Greenbelt, MD 20771
U.S.A

## SUMMARY

The Dynamics Explorer program was a dual spacecraft mission designed to study the interactions between the upper atmosphere, ionosphere and magnetosphere. The global auroral images acquired from the high altitude spacecraft have provided a revolutionary new time-dependent frame of reference for the interpretation of data acquired in-situ. Using data especially from the low altitude spacecraft, interrelationships are developed between the various electrodynamic parameters measured. Ionospheric irregularities are found to be especially intense in regions of electric field convection shears, which are closely related to the dusk hemisphere field-aligned currents. These region 1 currents are spatially connected to the boundary plasma sheet electron precipitation. At all local times, there appears to be a universal relationship between regions where div E < 0 and electron precipitation structures. Depending upon the characteristics of the electrons bombarding the atmosphere, the atmosphere will radiate various spectral optical emissions, which can be imaged from high above the polar caps, with temporal resolution sufficient to follow the time-dependent evolution of a substorm. It is concluded that with further detailed analyses of the electrodynamic parameters obtained from in-situ measurements, analyses of auroral images will yield specific information on many of the important ionospheric parameters over an entire auroral oval and polar cap, including regions of intense ionospheric irregularities.

## INTRODUCTION

The theme of this meeting declares that "much of the local ionospheric behavior is determined by coupling electrically to the distant ionosphere and solar wind, and coupling electrodynamically to higher and lower levels of the atmosphere". NASA's Dynamics Explorer program was designed specifically to study this electrodynamic coupling between the upper atmosphere, ionosphere and magnetosphere. The theme proceeds to point out that global observations are often necessary to provide the frame of reference for interpreting observed effects. Indeed the Dynamics Explorer program acquired for the first time observations of a global nature which provide a revolutionary new time-dependent frame of reference for the interpretation of locally observed phenomena. Therefore, using data acquired from the Dynamics Explorer mission, this paper will strive to show how the various measurements of the electrodynamic parameters in near-Earth space are related, and will begin to place them into this global frame of reference. Only by developing such relationships within a time-dependent framework can time-dependent, and, therefore, realistic models of the ionospheric behavior be eventually developed.

Much of the material in this review has been taken from a set of review articles written by members of the Dynamics Explorer Science Team, which will appear in Reviews of Geophysics (Table 1). While this review is limited predominantly to auroral and polar cap phenomena, these phenomena also display effects all the way to the equator.

The Dynamics Explorer program utilized two spacecraft launched together into coplanar polar orbits (inclinations of 90°). One spacecraft, known as DE 1, was placed into an elliptical orbit with an apogee of 3.63 $R_E$ altitude, initially over the northern polar cap. The second spacecraft, DE 2, was in a more circular orbit in the ionosphere, with perigee at about 300 km and apogee initially at about 1000 km. The spacecraft, mission and instruments have been described in an issue of Space Science Instrumentation [1]. Joint operations with the two spacecraft commenced shortly after the launch on August 3, 1981, and extended to February 19, 1983, when the DE 2 spacecraft reentered. DE 1 operations continue and are planned into the 1990s.

Table 1

| | |
|---|---|
| R. A. Hoffman | The Magnetosphere, Ionosphere, Atmosphere as a System: Dynamics Explorer Five Years Later |
| J. L. Burch | Energetic Particles and Currents: Results from Dynamics Explorer |
| Charles R. Chappell | The Terrestrial Plasma Source—A New Perspective in Solar Terrestrial Processes from Dynamics Explorer |
| L. A. Frank and J. D. Craven | Imaging Results from Dynamics Explorer 1 |
| D. A. Gurnett and U. S. Inan | Plasma Wave Observations with the Dynamics Explorer 1 Spacecraft |
| R. A. Heelis | Studies of Ionospheric Plasma and Electrodynamics and Their Applications to Ionosphere-Magnetosphere Coupling |
| T. L. Killeen and R. G. Roble | Thermosphere Dynamics: Contributions from the First Five Years of the Dynamics Explorer Program |

## TIME-DEPENDENT FRAME OF REFERENCE

Well before the beginning of the space age, experimenters attempted to place their observations within the context of more globally occurring phenomena. For many years, and even today, the magnetic indices of Kp, the various A indices, and Dst have been used to describe the global activity of the earth's magnetic field. With the need for a higher resolution index to characterize polar disturbances, the various auroral indices of AU, AL and AE were developed. The problem with using these ground-based magnetic indices is often two-fold: first, the measurements or the calculational procedures upon which the indices are based do not have sufficient spatial or provide sufficient temporal resolution to properly characterize the disturbances; second, the use of the indices often removes them from their temporal context. Thus measured parameters are analyzed together as a function of the absolute value of the index, irrespective of whether the index shows an increasing or decreasing tendency.

In addition to the magnetic indices, all-sky-camera photos have been used to place local measurements into the context of at least the local sequence of auroral activity.

We now have available a new tool, "global auroral images", which are, as the name implies, global in their coverage, at least to an entire auroral oval and polar cap, and are sequential in their acquisition [2]. By relating in-situ measurements of electrodynamic parameters to each other and to auroral emissions, a time-dependent understanding of the in-situ measurements on a "global" basis can be acquired.

The first image (Figure 1) compares the regions of coverage of a global auroral image in UV wavelengths (and, therefore, in false color) with that of an all-sky-camera, shown as the white circle [3]. The global image places the ground-based measurement, which has greater local temporal and spatial resolution but very limited spatial coverage, into the proper context of the instantaneous configuration of the dynamic auroral oval.

The figure also shows the distribution of the magnetic observatories from whose data the AE(12) index is determined [4]. Note that the array of stations was not properly located for considerable longitudinal spans or in latitude to provide information on the local time-dependence of the activity. In this case the dominant auroral enhancement near midnight lies between the stations of Tixie Bay and Cape Wellen.

The second figure contains a series of images taken at 12 minute intervals in which a modest substorm was viewed from a portion of the orbit above the northern hemisphere. These images, also in UV, demonstrate first the success of the instrumentation for viewing the entire auroral oval, even well into the sunlit portion of the earth. They also dramatically illustrate the dynamic nature of a substorm and the explosive release of energy. The onset of the substorm is evident in the fourth image and involved only a small area. A westward traveling surge then travelled along the poleward edge of the oval, as highly structured, eastward-moving auroral forms developed in the post-midnight sector. Note the large increase in the area of the active region. It is within such a dynamically changing framework that it is now possible to analyze smaller scale, in-situ measurements. No index, or statistically averaged oval can provide such a dynamic frame of reference.

In addition to a reference frame, such images can be interpreted to specify characteristic energies and energy fluxes of the electrons producing the optical emissions observed by auroral images [5]. Two spectral emissions are required to specify these quantities, such as those obtained from an ultraviolet channel and a visible wavelength channel. The ionospheric conductivities can then be calculated from the electron characteristics.

## CONVECTION PATTERNS

We will now begin to show relationships between the geoelectric field, magnetic field perturbations from field-aligned currents, electron precipitation characteristics and ionospheric irregularities which can be placed eventually into the temporally and spatially varying global auroral framework. We begin with large scale convection patterns.

The most extensive development of empirical models of the convection electric field has been carried out by Heppner and Maynard [6], using measurements of a single horizontal component of the electric field acquired from the DE 2 satellite [7]. Patterns were derived in relation to various directions of the interplanetary magnetic field (IMF). Figure 3 contains two ionospheric convection patterns derived from the electric field measurements for the northern polar region when the IMF was predominantly southward ($Bz < 0$), one for $By > 0$ (duskward), the other for $By < 0$. The two patterns differ in several characteristics:

1. Most obvious are the relative sizes of the two convection cells, which reverse with the reversal in By.
2. A noon overlap of the two cells also reverses with By.
3. For $By > 0$, the dawn cell protrudes into the dusk cell in the pre-midnight hours. We will return to this point later.
4. The foci of the two cells are rotated clockwise 1 to 2 hours from dawn-dusk. These patterns have been substantially confirmed by radar [8, 9] and other satellite [10] measurements.

For these average patterns the potential between the foci of the two cells in both cases is 76 keV. It was found that the patterns do not change in shape with the level of activity, but the mean values of the polar cap potential drop show a magnetic index dependence, as seen in Figure 4. The average potential shown in Figure 3 corresponds to an AE index near 400, or a Kp between 3+ and 4-.

When the IMF has a substantial northward component ($Bz > 0$ or positive), a variety of convection signatures are observed in the usual polar cap. The interpretation of these signatures has led to two different constructions of large scale convection patterns, distorted two cell patterns vs. multiple cell patterns. Despite the differing interpretations there is general agreement that regions of

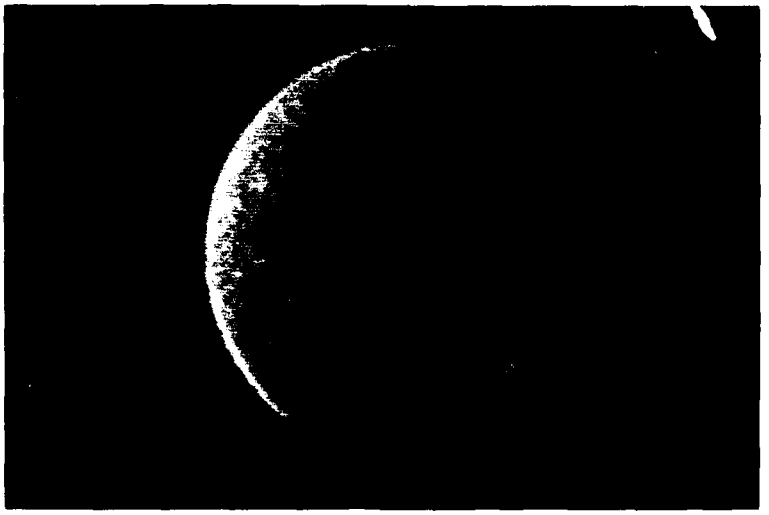


**Figure 1.** An auroral image of the northern polar region acquired from the Spin-Scan Auroral Imager on DE 1 [2] at 10:58 UT on November 23, 1981. This image shown in false color was obtained in 12 minutes at vacuum ultraviolet wavelengths 123 to 155 nm. The crescent shaped region of emissions on the left is the sunlit earth, whereas the large oval is the auroral oval. The locations are indicated of the twelve ground-based magnetic observatories which provide data for the determination of the AE auroral indices [4]. The small oval between the letters T and W shows the field-of-view of an all-sky-camera for 10° elevation angle to an altitude of 120 km centered at 63.3° N magnetic latitude at 21 hours MLT.



**Figure 2.** Twelve consecutive false-color images of the northern auroral oval recorded the onset and evolution of a substorm as the spacecraft Dynamics Explorer 1 descended from apogee during the period 0529 through 07:55 UT on 2 April 1982. Each image was telemetered in a twelve-minute interval. The sunlit atmosphere is visible at the upper left in each image. The passband of the optical filter is 123 to 155 nm, for which the dominant responses are from emissions of atomic oxygen at about 130.4 and 135.6 nm and the LBH bands of molecular nitrogen

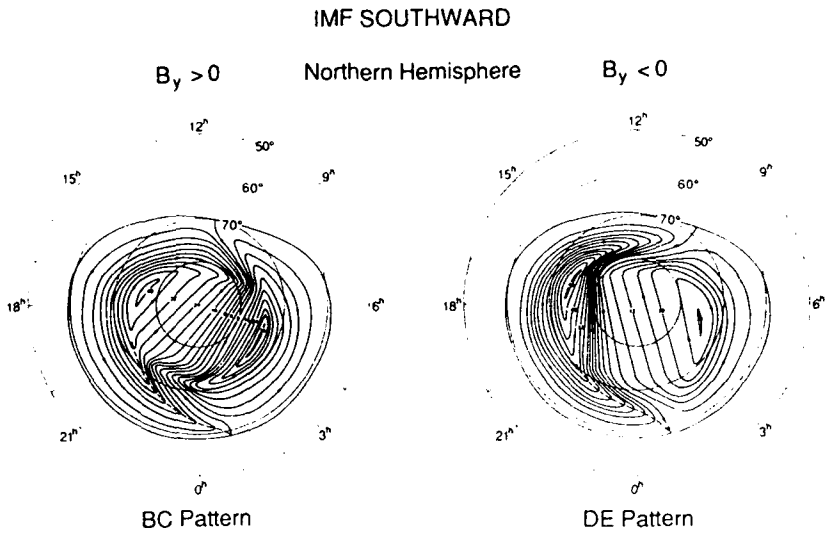



**Figure 3.** Convection electric field model BC (left), representing the most typical pattern distribution encountered in the northern hemisphere for +Y IMF directions and in the southern hemisphere for -Y IMF directions. Convection electric field model DE (right), one of two models representing the pattern distributions encountered in the northern hemisphere for -Y IMF directions (+Y IMF for southern hemisphere).

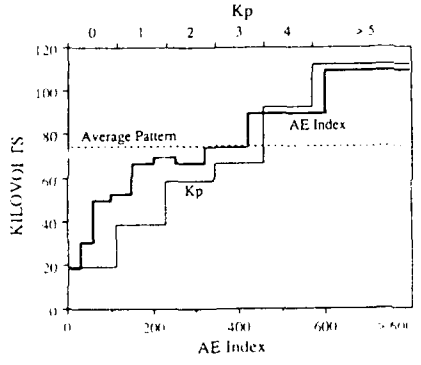



**Figure 4.** Mean values of the polar cap potential drop from both hemispheres and all signatures during November and December 1981 as a function of AE and Kp indices (from [6]).

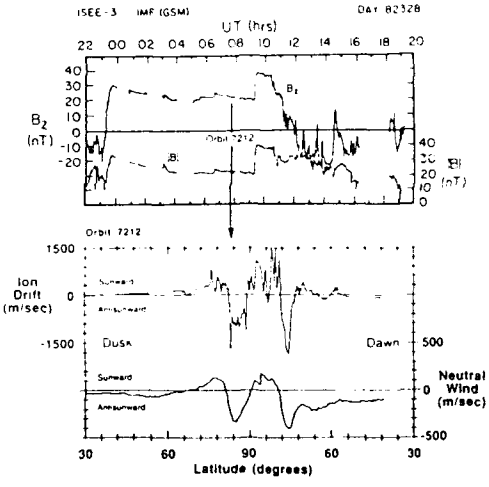



**Figure 5.** Variation in the Bz component and B magnitude of the interplanetary magnetic field as measured by ISEE 3 during 24 November 1982 (top). The time of a DE 2 pass (orbit 1712) is marked. The ion drifts and zonal neutral winds measured by DE 2 during the dawn to dusk orbit 1712 as a function of geographic latitude (bottom). Note the 3:1 scale difference (from [11]).

sunward convection exist at most times at high invariant latitudes where normally antisunward convection would occur if the IMF were southward. An important confirmation of stable sunward convection, rather than a transient condition, comes from a comparison of the bulk motions of both ionic and neutral species in the high-latitude F-region [11]. In Figure 5, top, the ISEE 3 data show a strongly northward, stable IMF for the first half of the day, during which time a sequence of DE 2 passes occurred. The zonal (sunward-antisunward) components of both the ion drift and neutral wind in the F-region from one of these passes are compared in the lower portion of the figure, with a 3:1 scale difference. It is very evident that throughout the polar region the zonal neutral wind follows the "averaged" zonal ion drift, and contains a well defined region of sunward flow in the central polar cap. Since the time constant for the ion gas to set the entire neutral gas in motion in the winter polar F-region is on the order of 2-3 hours, we conclude that sunward ion convection in the polar cap at times of northward IMF can provide a sufficiently strong and long-lasting momentum source for the neutral gas to overcome the effects of the antisunward directed pressure gradient force. Therefore, we consider the measured sunward ion convection to be a characteristic of positive and steady Bz conditions and not a transient flow regime.

How is the sunward convection interpreted in the context of a global pattern? In their development of empirical convection models, Heppner and Maynard [6] included this sunward flow by progressively distorting the Bz southward models as Bz becomes

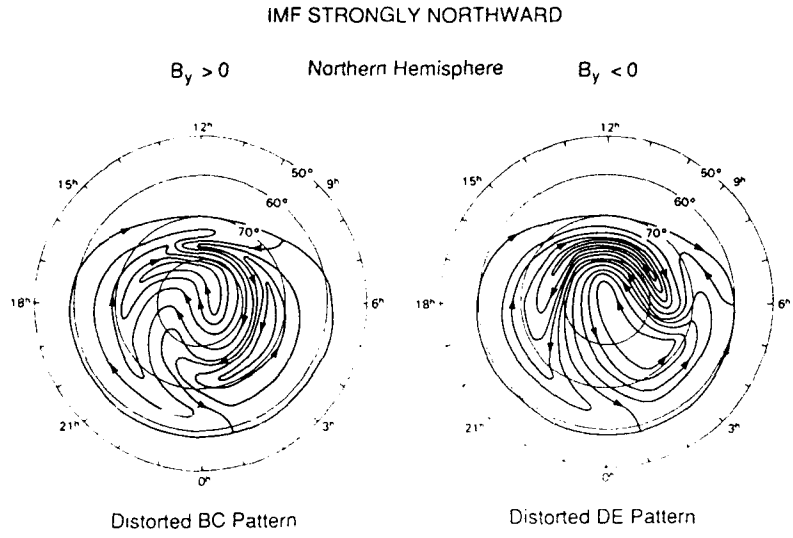



**Figure 6.** Rotationally twisted distortions of model BC (left) and model DE (right) usually observed when the IMF Bz is strongly northward (positive) (from [6]).

more strongly northward. For the case of Bz strongly northward, their By > 0 and By < 0 models appear in Figure 6. In both cases, the distortions have a common characteristic: the evening foci rotate toward noon and the pattern is stretched into the morning hours. This results in an inverted "U"-shaped convection pattern over the polar cap, containing a region of sunward flow.

On the other hand, Heelis et al. [12] connect equipotential flow lines into closed loops (Figure 7), resulting in a multicell convection pattern in the polar cap, as Burke et al. [13] had done previously. Heppner and Maynard [6] object to such closed loops on the nightside on the basis that sign reversal sequences in the midnight sector are essentially unchanged during Bz positive conditions from those observed for Bz negative.

Part of the problem with Bz (+) conditions arises from the general suppression of auroral and magnetic activity under this condition. In an extensive study of the electrodynamic patterns in the polar region during periods of extreme magnetic quiescence, Hoffman et al. [14] found that convection patterns were still likely to be found, though contracted and with large amplitude irregularities superposed. The probability of identification of any type of large scale pattern decreased significantly around the winter solstice, an example of which is shown in Figure 8. Note the many reversals in the electric field direction (second panel), each negative gradient associated with an electron precipitation structure, but no large scale pattern can be identified.

At this time, we have ample evidence that the geoelectric field pattern is highly variable and dependent upon the strengths of the various convection drivers arising from the interactions of the IMF and the geomagnetic field. However, even a crude time-dependent empirical geoelectric field model does not yet exist.

## FIELD-ALIGNED CURRENT PATTERNS

Global field-aligned current patterns have been constructed primarily on the basis of statistical averages. The extremely well known region 1 and region 2 patterns [15], which pertain to two levels of magnetic activity, have been refined especially with respect to the effects of the northward IMF (see [16] for a review).

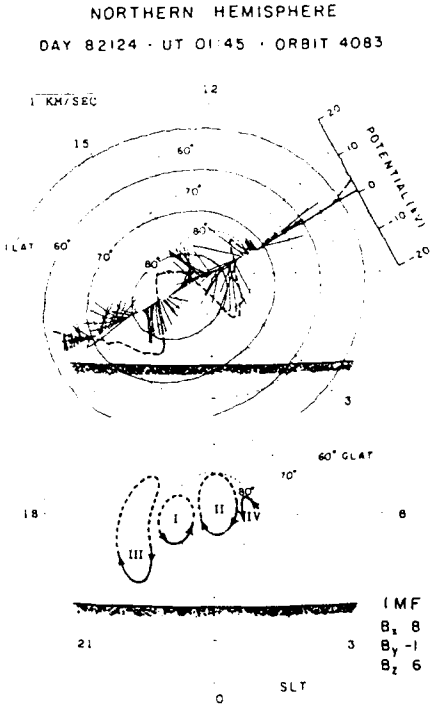



Figure 7. High-latitude convection signature observed when the IMF Bz is positive and By is small. Ion drift vectors and the associated electrostatic potential are shown in the top dial. A multicell interpretation of representative equipotential flow contours are shown below. The shaded arc denotes the terminator at 200 km (from [7]).

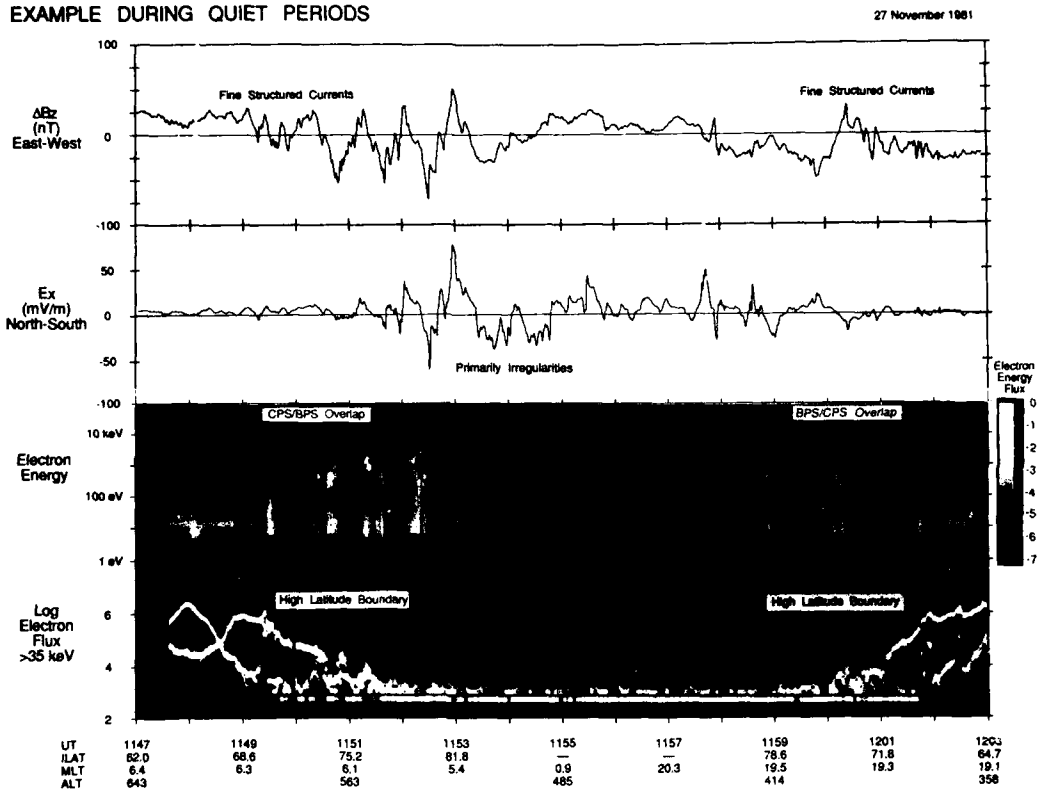



**Figure 8.** Data from DE 2 taken during a pass over the northern polar region at the time of extreme magnetic quiescence. The top panel contains the east-west component of the magnetic field after the field from the internal field model was subtracted. Positive values denote an eastward orientation. A positive gradient indicates a downward current during a poleward segment of the pass and an upward current during an equatorward segment. The second panel shows the north-south component of the electric field, with positive values in the direction of the spacecraft motion (antisunward convection). The third panel is an electron spectrogram for energies from 5 eV to 30 keV, with energy flux represented by a grey scale with units of ergs/cm$^2$-sec-ster-eV. The bottom panel contains plots of the intensities of electrons having energies greater than about 35 keV, with Xs designating the flux in the loss cone, and Os the trapped fluxes.

Unfortunately, the significant modifications to the patterns during substorms [17] have been largely ignored.

Recently we have been attempting to acquire an understanding of the relations between the field-aligned currents, convection patterns and particle precipitation patterns as a function of substorm phase [18, 19, 20] This work has been based to a large extent on the frequently prevailing coupling mode between the perturbation magnetic field, interpreted as due to field-aligned currents, and the electric field, expounded on especially by Sugiura [21]. This mode states that the field-aligned currents and electric field are orthogonal to each other, and are highly correlated. An example of this correlation is shown in Figure 9 with an extremely high correlation of 0.992. There are three important consequences of this proportionality:

1. $\nabla \cdot J_H = 0$ The height integrated Hall current is divergence free, or the field-aligned current closes in the ionosphere by Pederson currents.
2. $\Sigma_p = \frac{\Delta Bz}{\mu_0 Ex}$ The height integrated Pederson conductivity is obtainable from the ratio of the perturbation magnetic field and the electric field measured in the topside ionosphere.
3. $\Delta\vec{B} \parallel \vec{V}$ (plasma), since $\Delta\vec{B}$ and $\vec{E}$ are orthogonal. The direction of the perturbation magnetic field is parallel or antiparallel to the plasma convection. We will investigate these consequences in more detail later, but for now will utilize the proportionality feature to begin construction of global time-dependent or substorm phase dependent electrodynamic patterns.

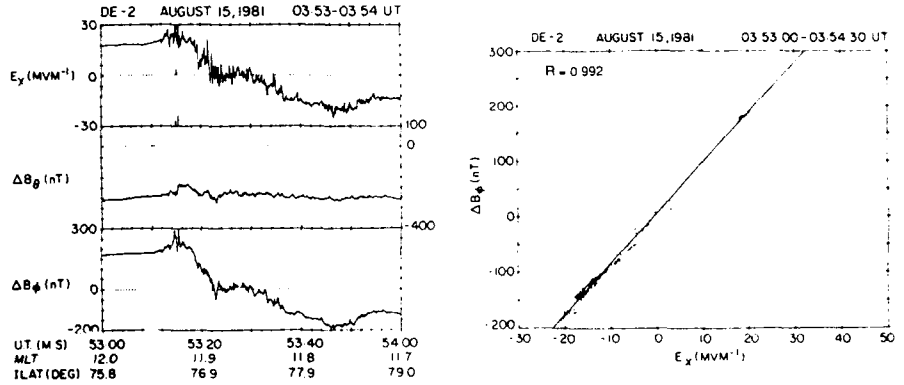



**Figure 9.** Data from a pass of DE 2 through the cusp (left). The top panel contains the north-south component of the electric field, the lower two panels the north-south and east-west components of the magnetic field in geomagnetic coordinates. Both sets of data contain 16 samples/second. The plot on the right shows the correlation between the electric field and nearly orthogonal component of the magnetic field, with a correlation coefficient of 0.992 (from [21]).

Figure 10 is an example from the dusk local time during the maximum phase of an isolated substorm when the IMF By was slightly positive. Recall that the Heppner and Maynard model of the convection pattern for By positive has a large dusk convection cell with weak gradients, but the dawn cell overlaps the dusk cell, causing a swirl in the convection pattern (Figure 3, left). Thus the polar cap electric field (first $1^1/_2$ minutes in the middle panel) has a fairly constant value, followed by an electric field convection spike due to the swirl. Note the excellent relations between the region 1 current identified from the east-west magnetic perturbation, the negative gradient in the north-south component of the electric field and the boundary plasma sheet (BPS) electron precipitation (in this case a large inverted-V), displayed as a grey-scale spectrogram. A similar relation exists between the region 2 current and the positive gradient of the electric field, but not with the central plasma sheet (CPS) precipitation, which barely existed. Also note that no current is associated with the convection field spike. It is observed, however, that if the By component is large, a current oppositely directed to region 1 is usually associated with the positive gradient of the convection spike.

The preliminary analysis of these interrelationships using DE 2 data from 126 polar passes during relatively isolated substorms was previously reported [19]. A summary of the results is shown schematically in Figure 11, with field-aligned current regions defined from the magnetic field measurements overlaid on slightly modified versions of the Heppner and Maynard convection patterns for Bz negative and the two By orientations. When By is positive the strongest gradients in the convection pattern appear in the morning hours, with a convection spike at the high latitude edge of the region 1 currents, often with an associated oppositely directed current. Such a current has also been reported from Viking and DMSP observations and identified as a "region 0" system [22]. The protrusion of the morning convection cell into the pre-midnight hours brings a convection gradient and, therefore, an extension of the morning region 1 current above the evening region 1 current, producing the overlap found in the statistical patterns [15]. When By is negative, the strongest gradients occur on the dusk side, with evidence for an extension of the morning region 1 current into the dusk hemisphere as a region 0 current. The heavy dashed line is the Harang discontinuity, which runs through the evening region 1 current. Recent analysis has shown that it should be drawn closer to the high latitude boundary of the current region (R. Fujii, private communication, 1988). Recalling the AE dependence on the polar cap potential (Figure 4), as a substorm subsides, the potential gradients subside, and, therefore, the strengths of the field-aligned currents lessen.

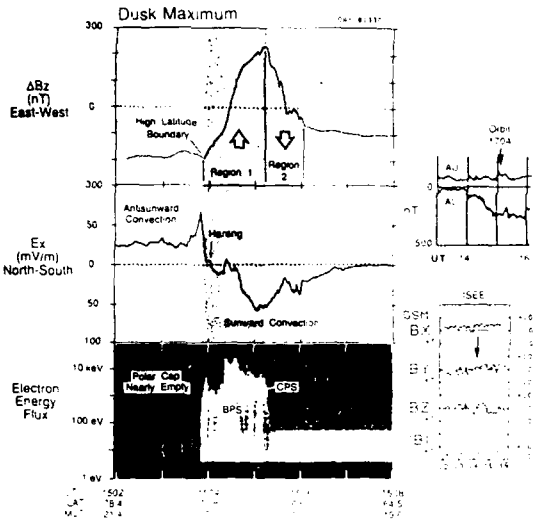



**Figure 10.** Data from a pass of DE 2 through the northern auroral oval at about 20 MLT. At the time of the pass the IMF had a negative Bz component and a slightly positive By component. A small isolated substorm had reached its maximum phase. The magnetic field perturbation in the east-west direction displayed a clear region 1 and region 2 current system (top panel). The north-south electric field component showed antisunward convection over the polar cap, a convection enhancement just prior to 15:04 UT, and a negative bay of sunward convection at lower latitudes (middle panel). The electron spectrogram for electrons with 15° pitch angle in the lower panel shows a polar cap nearly devoid of precipitation, a wide boundary plasma sheet (BPS) inverted-V, and an extremely narrow region of diffuse central plasma sheet precipitation.

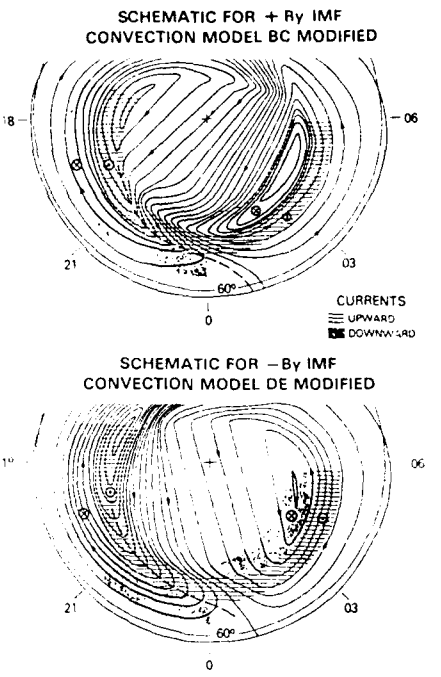



**Figure 11.** Schematic diagrams of the regions of upward and downward field-aligned currents overlaid on slightly modified versions of the convection patterns shown in Figure 3.

For periods of extreme magnetic and auroral quiescence, an identification of the region 1 and region 2 field-aligned current system disappears, except in the cusp region [14]. Even there they appear to be absent in mid-winter. Field-aligned currents continue to exist, but as fine structure distributed through the auroral oval and over the polar cap. An example of these fine-structured currents as detected by the magnetic field perturbations appears in the top panel of Figure 8. Note the excellent correlation with the electric field variations.

From this work we have seen how, at least to a first approximation, we can interrelate field-aligned current regions, convection patterns and part of the electron precipitation patterns into a self-consistent model, and associate the changing patterns at least with the IMF direction. In recent analyses, we are extending this work specifically to the phase and location relative to the substorm auroral enhancement identified by the auroral images acquired simultaneously from DE 1 [23]. Figure 12 contains an auroral image taken during the maximum phase of a substorm, and displays a bulge-type auroral expansion. The DE 2 spacecraft transited the bulge, and the data appear in Figure 13. An upward field-aligned current is observed in the poleward portion of the bulge and seems to be separated into two components: a strong higher latitude current associated with an electric field characteristic of the polar cap (antisunward), and a lower latitude current associated with the auroral oval, or sunward convection region. The electron precipitation pattern can also be interpreted as containing two regions of inverted-Vs. Downward currents are observed poleward and equatorward of the upward current, the high latitude component associated with the convection spike. Multiple current reversals often exist within the bulge region, evidenced by the multiple sign changes in the gradient of the magnetic field between 10:20 and 10:21 UT.

An analysis of 27 passes at different locations with respect to the bulge has produced the pattern of currents shown in Figure 14. The poleward downward current and convection spike

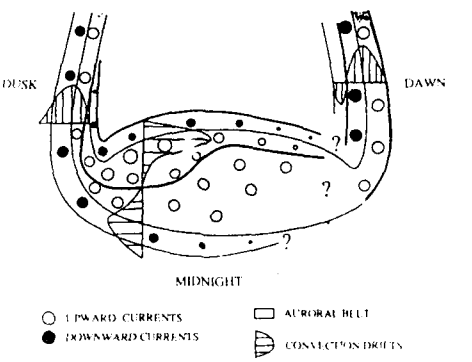



**Figure 14.** A schematic diagram of the large-scale field-aligned currents and convection pattern in a bulge- type auroral enhancement. The heavy line through the auroral belt is the Harang discontinuity.

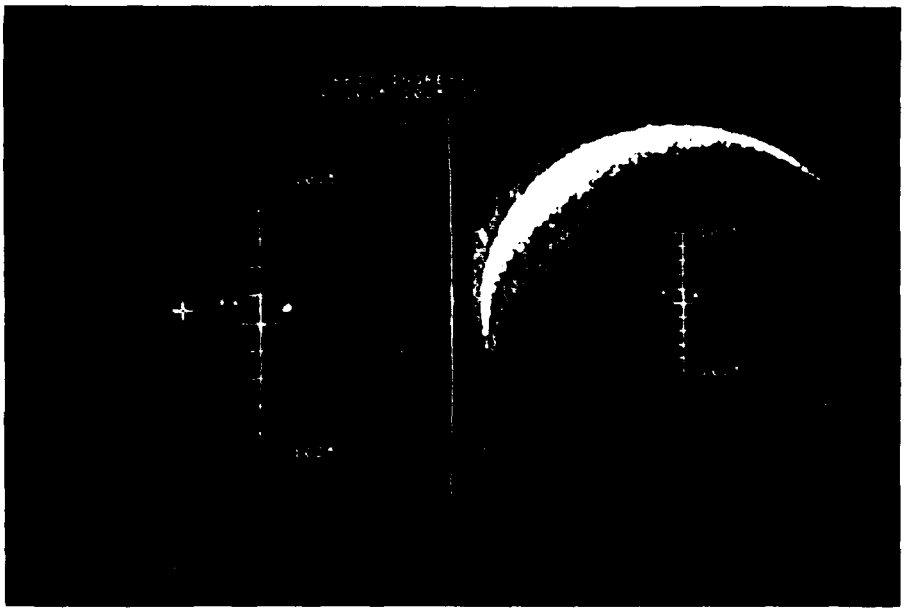


**Figure 12.** Auroral image of a bulge-type auroral enhancement in the midnight sector of the northern auroral oval on September 27, 1981, with the ground track of DE 2 superposed (right). The bulge region is enlarged in the left panel and shows a bright arc along the poleward edge through which the DE 2 satellite passed.

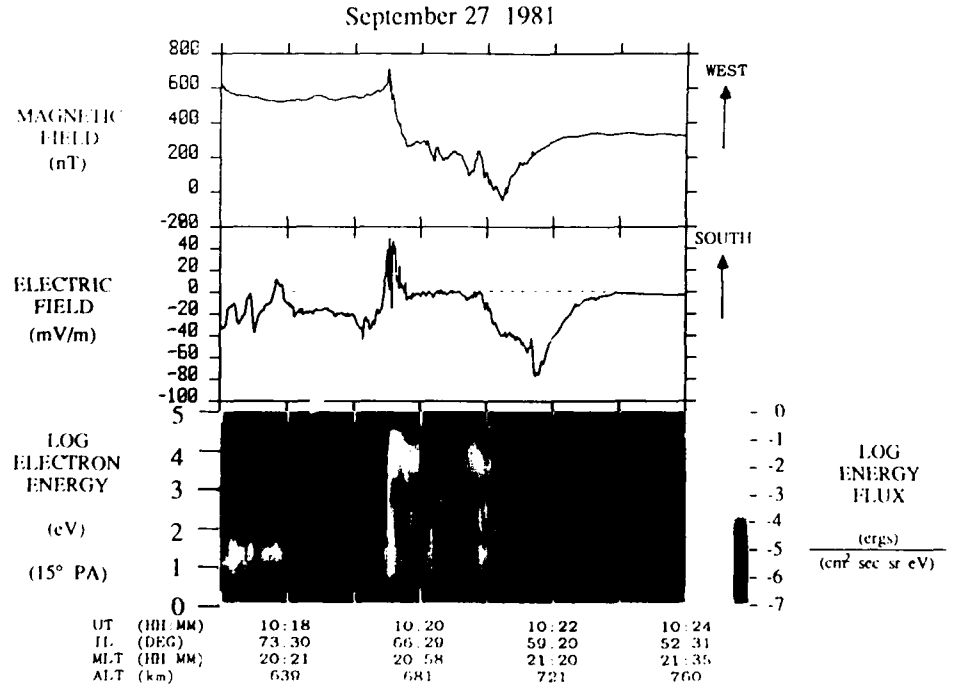



**Figure 13.** Data from the DE 2 pass through the auroral bulge shown in Figure 12. The top panel contains the east-west component of the magnetic field perturbation. A negative gradient indicates an upward field- aligned current. The middle panel shows the north-south component of the electric field. The bottom panel is a grey scale spectrogram of electron precipitation. The inverted-V electron structure from about 10:19:30 to 10:19:50 UT is coincident with the bright arc along the poleward edge of the bulge region.

become smaller and disappear both westward and eastward of the bulge, leaving only the usual region 1 and region 2 currents. The heavy line through the auroral belt indicates the Harang discontinuity, the division between counter-clockwise (high latitude) and clockwise (lower latitude) convection. On the dusk side of the bulge, nearly all currents lie within the sunward convection. The wide region of average upward currents within the bulge usually contains multiple currents of both directions. Currently we are questioning whether the region 1 dusk-side current is connected continuously with the dawn-side region 2 current, as shown in the statistical studies [17] and in Figure 11. The details of the convection and current patterns in this early morning portion of the bulge are not clear.

## FIELD-ALIGNED POTENTIALS

Before investigating in detail the relationships between the electrodynamic parameters in auroral precipitation structures, the subject of electric potentials along magnetic field lines will be discussed. While the existence of some type of potential structure along magnetic field lines has been generally accepted, it has not been based on much definitive evidence. These potentials are a form of decoupling between the magnetosphere and ionosphere. The two most important studies performed with data from the DE spacecraft made use of the unique orbital configuration of the two spacecraft: the high and low apogee coplanar orbits, which provide simultaneous measurements of aurorally related parameters at two altitudes.

Reiff et al. [24] evaluated the velocity distributions of ions and electrons acquired nearly simultaneously at two altitudes approximately along the same magnetic field lines to test for the effects of any electric fields aligned parallel to the magnetic field. They used three techniques to estimate the potential drop between the two spacecraft when DE 1 was between 9,000 and 15,000 km and DE 2 was between 400 and 800 km:

1. The difference between the peak electron energies in the inverted-V electron structure.
2. The peak energy of upward flowing ions at DE 1. No upward flowing ions were observed at DE 2.
3. The enlargement of the atmospheric loss cone formed by upgoing electrons as they are decelerated by the potential drop below DE 1, but above DE 2.

Of course, all techniques should provide the same values for the potential should all effects be purely electrostatic. Somewhat surprisingly, the peak ion beam energy was frequently 30% (and in one case 70%) less than the potential drop below DE 1 as derived from the electron distributions. The authors did find, however, that the average energy of the ions moving parallel to the field agree better with the electron results. An example comparison of the potentials obtained by the two methods using electrons, and the average ion energy is shown in Figure 15. DE 1, at 2.7 $R_E$ was above or nearly above the acceleration region, so $e\Phi_{IV} = e\Phi_{LC}$. The fact that all three, quite independent techniques yield evidence for a similar shaped potential structure of comparable magnitude is possibly the best proof to date of the existence of at least a quasi-static potential structure. The fact that the average parallel ion energy rather than the peak energy correlated best with the energies derived by the two electron measurements has led to the conclusion that an ion heating process is involved within the auroral acceleration region.

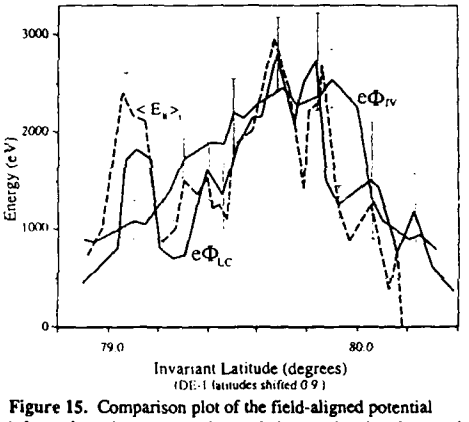



Figure 15. Comparison plot of the field-aligned potential inferred from three independent techniques using data from both DE 1 and DE 2 during a magnetic conjunction: $e\Phi_{IV}$, from the peak energies in the inverted-V electron precipitation spectra measured by DE 2; $e\Phi_{LC}$, from the widening of the loss cone of electrons measured by DE 1; $<E>_\parallel$, from the average parallel energy of upflowing ions measured by DE 1 (from [24]).

Again using nearly simultaneous measurements at two altitudes by DE 1 and DE 2, Weimer et al. [25] have examined the mapping of small scale structures in the electric field along magnetic field lines. They plotted the measured electric fields, which are usually perpendicular to the magnetic field lines, projected to a common altitude (Figure 16), and found that the large scale electric field is the same at both altitudes. Superposed on the large-scale fields, however, are small-scale features with wavelengths less than 100 km which are larger in magnitude at the higher altitude. Electric field spectra from a Fourier transform of the data show that the magnitude differences depend upon wavelength (Figure 17). The small scale variations are associated with field-aligned currents which are measured. The experimental measurements are found to agree with a steady state theory which postulates that there are parallel potential drops associated with the variations in the perpendicular electric fields. Furthermore, Fourier transforms of the east-west magnetic component measured by DE 1 and the north-south, or orthogonal, component of the electric field at DE 2 altitudes are found to be nearly identical, with the ratio giving the height integrated Pederson conductivity. The conductivities so derived are in good agreement with the conductivities calculated from the electron precipitation measurements.

These two separate types of analyses, one using charged particle data, the other electric field data, provide rather convincing proof for the existence of at least quasi-static electric field potentials along lines of magnetic force on scales less than about 100 km, or 1°, in latitude.

## ELECTRODYNAMICS OF ELECTRON PRECIPITATION STRUCTURES

We next investigate the details of the electrodynamic relationships in electron precipitation structures. In preparing the paper on field-aligned currents as a function of substorm phase for the IAGA meeting [19], we noticed that consistently the electric field showed

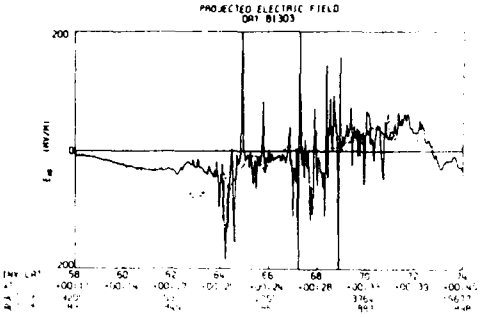



**Figure 16.** Electric fields measured by DE 1 and DE 2 projected to a common altitude for near simultaneous passes through the evening auroral oval (DE 1 MLT about 20.65, DE 2 MLT about 20.9). ΔT is the time DE 2 trailed DE 1 through the oval (from [25]).

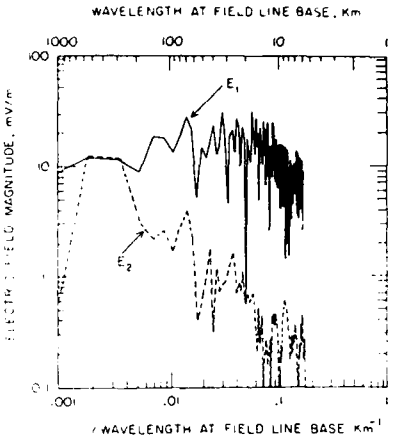



**Figure 17.** Electric field spectra from data displayed in Figure 16. The spectra are obtained from Fourier transforms of the electric field data between 65° and 70° invariant latitude. The ordinate values are the square root of the "spectral power density". The actual units are (mV/m) km (from [25]).

a gradient in the region of the electron precipitation, or upward field-aligned current, independent of the scale size of the current region, local time, or phase of the substorm. Of course, this is consistent with Sugiura's frequently prevailing coupling mode between the perturbation magnetic field and the electric field discussed earlier. Examples of two greatly different scale size events are seen by comparing Figures 10 and 18. The case in Figure 10 has a size of about 250 km and comprises the entire evening region 1 current. The two structures marked in Figure 18 extend over about 50 and 110 km, and exist in a sequence at usual polar cap latitudes. The unmarked negative gradient at 11:58:35 UT has a width one tenth the width of the inverted-V in Figure 10. Many examples of this pattern can be seen in Figure 8 also. Note, however, that for the case in Figure 10, the electric field tends to a value near zero at the time of the most intense electron precipitation (15:04:30 UT). Such an effect has been reported a number to times, especially from sounding rocket observations [26, 27, 28]. We will investigate this effect in more detail, since it appears to deviate from the correlation between the magnetic field perturbation and the electric field.

Data from an event similar to that in Figure 10 are plotted in Figure 19. Again the entire region 1 current (positive gradient in Bz) coincides with the wide inverted-V, which displays three maxima in energy. During the entire inverted-V the electric field is shorted to near zero. A convection spike occurred at the high latitude edge of the event, accompanied by a weak downward current. At the low latitude edge of the shorted region, the electric field plunged negative, and recovered correlated with the region 2 current (negative gradient in Bz), but not with the diffuse central plasma sheet (CPS) precipitation, which barely existed. By evaluating both the east-west and north-south components of the perturbation magnetic field, this event was found to be very nearly aligned to the geographic latitude, and, therefore, spacecraft coordinates (recalling that the orbit of DE 2 had a 90° inclination), and could be considered infinite in extent.

We investigate the applicability of the coupling mode between ΔB and E in Figure 20, which instead of a scatter plot of the correlation as shown in Figure 9, half second averaged values of the field components are plotted and connected with lines to follow the correlation in latitude across the event. The plot covers the entire region where the gradient of E is negative, including the shorted region, or the region 1 current. The correlation plot has three segments, the initial drop in E, the shorted region where the magnetic field continues to decrease, and the final plunge of the electric field. The heavy straight line is a least-squares fit to the data, and has a correlation coefficient of 0.59. It is quite apparent that this correlation is meaningless, and while

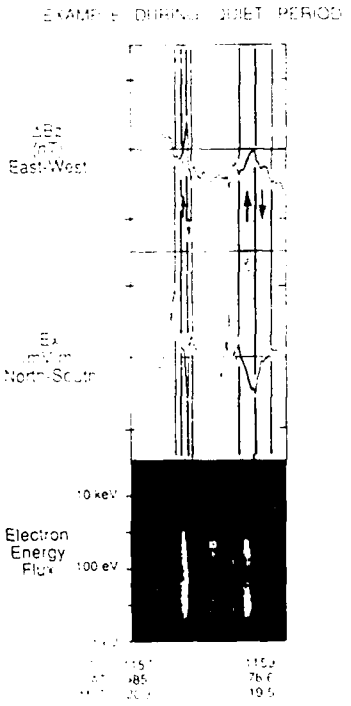



**Figure 18.** Together with Figure 10, examples of negative electric field gradients with various scale sizes coincident with upward field-aligned currents and electron precipitation structures.

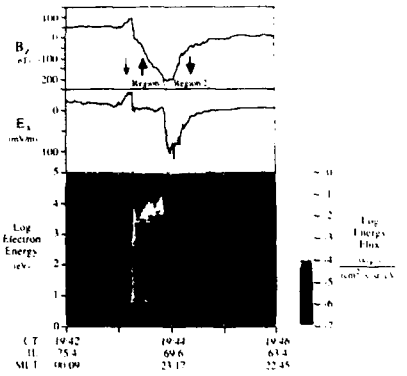


**Figure 19.** Example of electric field shorting in a region of strong electron precipitation near magnetic midnight. An energetic inverted-V electron precipitation structure (bottom panel) between 19:43 and 19:44 UT has maximum energies of about 10 keV. Coincident with this electron precipitation the ionospheric electric field (north-south component, center panel) shows a value near zero, whereas the negative gradient in the east-west component of the magnetic field indicates a region of upward field-aligned current (region 1).

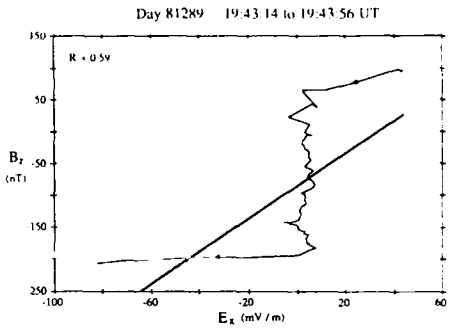



**Figure 20.** Correlation plot between the east-west component of the perturbation magnetic field and the north-south (orthogonal) component of the electric field for the region 1 upward field-aligned current region in Figure 19. A least-squares fit to the data is shown by the heavy straight line, which has a correlation of 0.59.

correlations to each segment individually would be very high, the ratio of the two fields (slope of the fitted curves) would not yield the height integrated Pederson conductivity. We find that the correlation between $\Delta B$ and $E$ holds only where the Pederson conductivity is uniform, i.e., where electron precipitation is of sufficiently low intensity and/or energy that ionization from the electrons is relatively weak, and, therefore, does not modify the conductivity.

To investigate this type of nighttime event further, we plot in Figure 21 (upper panel) the two components of the ion drift. The north-south component remains relatively steady across the entire event, with a weak equatorward drift. The east-west component (corresponding to the north-south component of the electric field) again shows the high latitude convection spike, then a strong westward drift equatorward of the electron precipitation region. The lower panel contains three calculations of electric potentials, the first from the ionospheric electric field, $\Phi_E = -\int E(N\text{-}S)\cdot dl$ (with zero potential at the low latitude side of the event). The potential increased with the convection spike, remained constant over the shorted region, and rapidly dropped towards zero in the region 2 current. This would be the potential in the magnetosphere also if the magnetic field lines were always equipotentials. However, interpreting the inverted-V event as measures of the field-aligned potential we need to add this potential ($\Phi_{TV}$) to the ionospheric potential to obtain the magnetospheric potential $\Phi_{total}$, as shown in Figure 21. This potential has extreme gradients or high electric fields at both edges of the inverted-V pointing into the event (div E < 0), and within the event, the electric field would show three regions of div E being negative, corresponding to the three parts of the inverted-V. It is such magnetospheric regions, which have negative space charge, that according to Lyons [29, 30] are the sources of upward field-aligned currents, or electron precipitation.

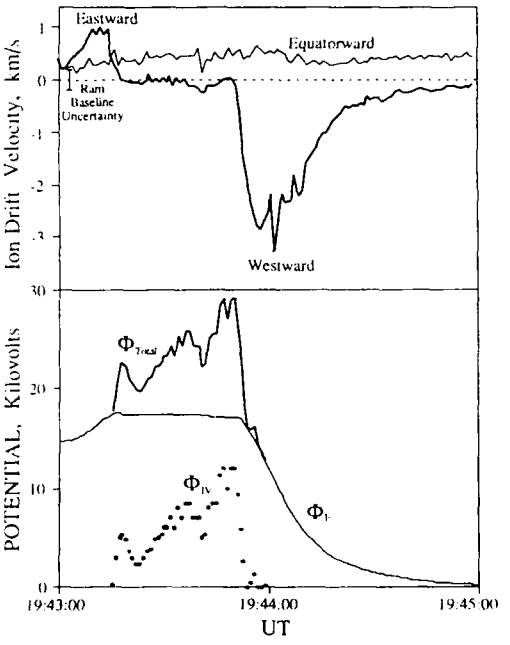



**Figure 21.** Two components of the ion convection drift (top panel) for the event in Figure 19. Three electric potentials calculated from the electric field and electron precipitation measurements (bottom panel): $\Phi_E = -\int E(N\text{-}S)\cdot dl$, where E is from Figure 19 and the potential is set to zero at latitudes below the auroral oval; $\Phi_{TV}$ is the peak energy in the inverted-V electron precipitation spectrum in Figure 19, which assumes the electrons were accelerated by an electric field along the magnetic fields; $\Phi_{total}$ is the sum of these potentials and would be the potential in the magnetosphere.

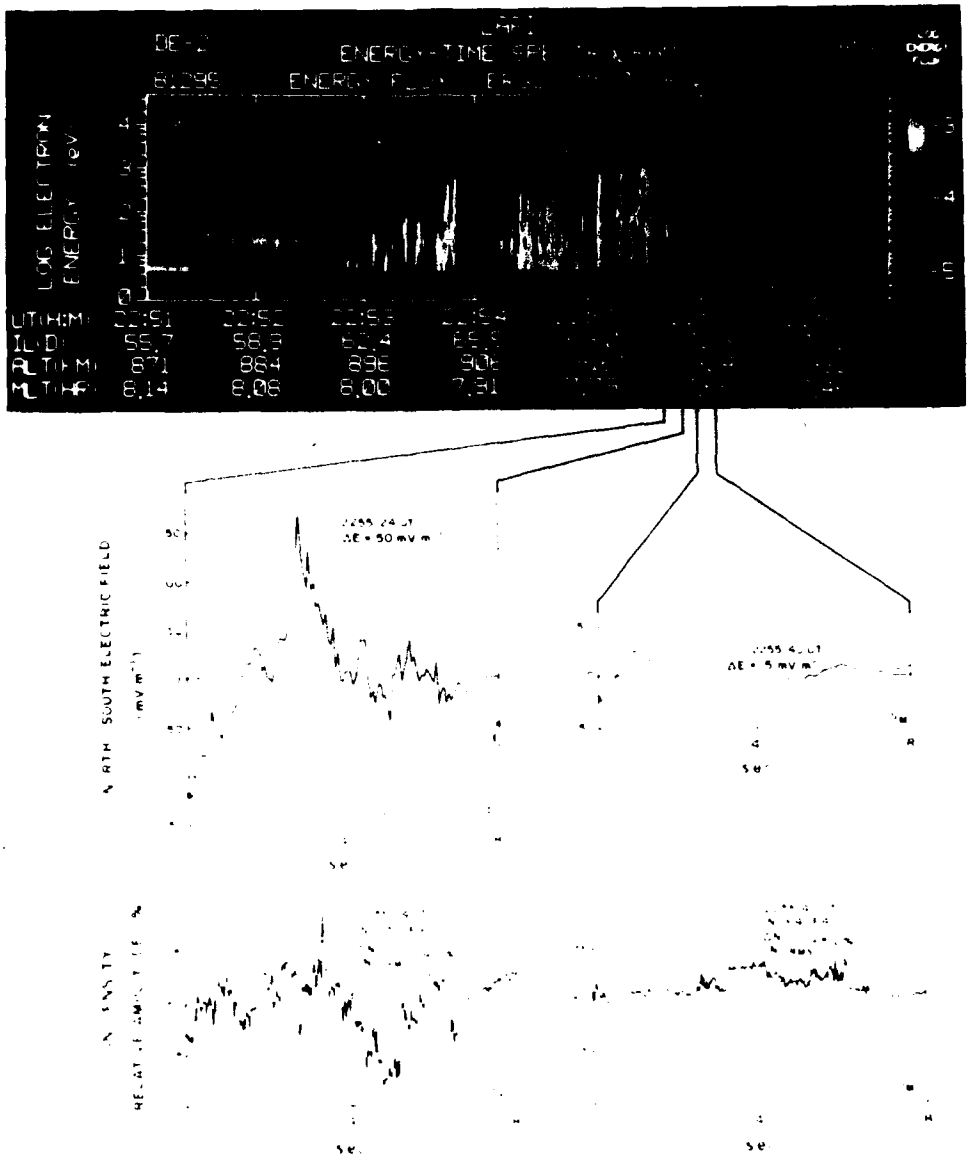



**Figure 23.** Electron spectrogram from DE 2 obtained from a pass through the morning auroral oval. The uniform precipitation between 100 eV and 1000 eV around 22:52 UT is the diffuse precipitation (mantle auroral precipitation [32]) with patches of higher energy diffuse precipitation extending to 22:54:30 UT. Beginning about 22:53 UT, electrons with energies below 1000 eV display characteristic structure, and end at 22:55:30 UT with the most intense event. Polar rain then extends over the polar cap. The middle panels each contain 8 seconds of electric field data (north-south component). The first case includes the time of the intense electron precipitation event and shows intense fluctuations and a large shear (negative gradient); the second case taken in the polar cap shows more modest fluctuations. The corresponding density fluctuations appear in the lower panels (from [31])

A summary of the ionospheric electrodynamics for this type of late evening auroral oval crossing appears in Figure 22. The satellite emerged from the polar cap where the electric field was weak with convection predominantly in an eastward direction (DE convection pattern, Figure 3). It encountered the convection spike, but the direction was unchanged. Upon entering region 1, the convection speed went nearly to zero, but changed direction to be primarily equatorward. At the boundary between region 1 and region 2, the convection speed rapidly increased, and the direction changed to westward. Thus the negative divergence in the electric field produces a convection reversal (Harang discontinuity), but the location is undefined within region 1.

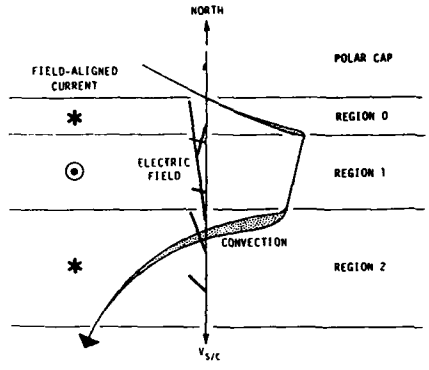



**Figure 22.** Schematic diagram of the convection through the electric field shorted event of Figure 19. The width of the convection line is a qualitative indicator of the speed of convection. On either side of the region 1 current, which is coincident with the energetic electron precipitation, the convection is predominantly parallel to the quasi-infinite (east-west) current sheet, and of opposite directions.

## IONOSPHERIC DENSITY AND ELECTRIC FIELD IRREGULARITIES

Associated with the velocity shear regions just discussed, the DE 2 measurements show considerably enhanced fluctuations in both the ionospheric F region and topside ionospheric ion densities and electric field. Basu et al. [31] analyzed data from two dawn-side passes and compared fluctuation spectra in the regions of high velocity and modest velocity shears. The example in Figure 23 contains an electron spectrogram taken in the morning auroral oval. An intense velocity shear is associated with the last significant electron structure at 22:55:30 UT (middle left panel), with the electron precipitation commencing with the maximum positive electric field and continuing during the negative gradient. On the other hand, the fluctuations in the electric field are much more modest in the polar cap (middle right panel), which was filled with polar rain. The corresponding ion densities are shown in the lower panels, with the RMS fluctuations almost three times larger in the high shear region compared to the polar cap. Density fluctuation spectra and electric field spectra were estimated by the maximum entropy method [31] for periods during the two levels of electric field fluctuations. A representation of the types of spectra found are shown in Figure 24. The observed spectral indices for both the $\Delta N/N$ and $\Delta E$ spectra lie around -1.8 down to wavelengths of a few hundred meters in regions of large velocity or electric field shears. At shorter wavelengths the electric field spectrum falls more rapidly, similar to the spectrum during moderate shear regions. While the ion density fluctuation spectra appear unchanged between the two cases, the power spectral density is one to two orders of magnitude larger in the large shear regions than in the surrounding regions studied.

## CONCLUSIONS

In this review from the Dynamics Explorer program pertaining to the electrodynamic coupling between the upper atmosphere, ionosphere and magnetosphere, we have shown interrelations between the various quantities measured, beginning with global empirical models of convection and working down to irregularities with scale sizes of a fraction of a kilometer. If this sequence is reversed, we

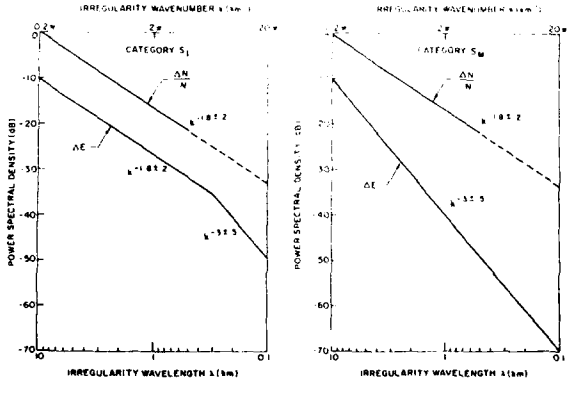



**Figure 24.** Idealized representation of the two categories of $\Delta N/N$ and $\Delta E$ spectra observed in association with velocity shears in the auroral oval: category $S_I$ for intense shears (~10 Hz), and category $S_M$ for moderate shears (~1 Hz). The $\Delta N/N$ and $\Delta E$ spectra have been arbitrarily separated. The short scale length ends of the $\Delta N/N$ spectra are shown dashed because of uncertainty in power spectral density estimates. (From [29]).

can see how the global patterns of ionospheric irregularities and other ionospheric phenomena might be monitored, eventually in near real time, by observing the auroral oval and polar cap with high resolution auroral imagers.

Ionospheric irregularities are especially intense in regions of electric field convection shears, which are closely related to the dusk hemisphere field-aligned currents (Figure 25). At all local times, there appears to be a universal relationship between regions where div E < 0 and electron precipitation structures. Depending upon the characteristics of the electrons bombarding the atmosphere, the atmosphere will radiate various spectral optical emissions, which can be imaged from high above the polar caps, with temporal resolution sufficient to follow the time-dependent evolution of a substorm. An analysis of the images taken simultaneously at two or more wavelengths can yield specific information on characteristics of the electron precipitation, from which ionospheric conductivities can be calculated. From relationships being developed between field-aligned currents as a function of substorm phase and MLT, some information on the location of region 2 field-aligned currents will be obtainable. Working through the characteristics of the electron precipitation, the location of velocity shears will be identified, which provides information on the time- dependent global pattern of ionospheric irregularities.

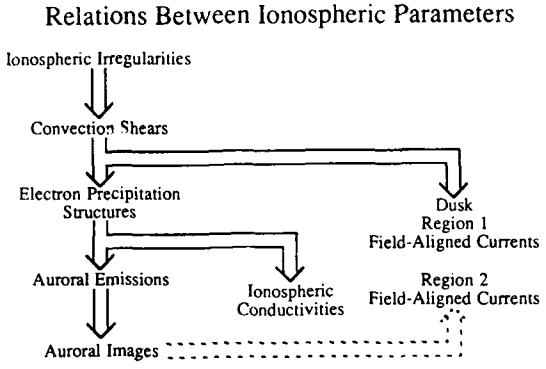



**Figure 25.** Relations between ionospheric parameters which indicate that regions of ionospheric irregularities can be related to regions of auroral emissions, which can be imaged globally from above the polar caps.

Admittedly much work remains to allow the details of the relationships to be defined over the myriad of conditions which the magnetosphere-ionosphere coupling processes present to us. However, the data bases exist between the Dynamics Explorer, HILAT, VIKING, radar and other programs. What remains is many years of analysis with the resource support for this analysis.

## REFERENCES

1. Hoffman, R. A., editor, Dynamics Explorer, *Space Sci. Instrum.*, **5**, 1981, pp. 345-561.

2. Frank, L. A., J. D. Craven, K. L. Ackerson, M. R. English, R. H. Eather and R. L. Carovillano, Global auroral imaging instrumentation for the Dynamics Explorer mission, *Space Sci. Instrum.*, **5**, 1981, pp. 369-393.

3. Frank, L. A., and J. D. Craven, Imaging results from Dynamics Explorer, sub. *Rev. Geophys.*, 1988.

4. Kamei, Toyohisa, and Tohru Araki, Data Book No. 10, Data Analysis Center for Geomagnetism and Spacemagnetism, Kyoto University, Kyoto 606, Japan, Nov. 1984, p. 2.

5. Rees, M. H., D. Lummerzheim, R. G. Roble, J. D. Winningham, J. D. Craven and L. A. Frank, Auroral energy deposition rate, characteristic electron energy and ionospheric parameters derived from Dynamics Explorer images, sub. *J. Geophys. Res.* Jan. 1988.

6. Heppner, J. P., and N. C. Maynard, Empirical high latitude electric field models, *J. Geophys. Res.*, **92**, 1987, pp. 4467-4489.

7. Maynard, N. C., E. A. Bielecki and H. F. Burdick, Instrumentation for vector electric field measurements from DE-B, *Space Sci. Instrum.*, **5**, 1981, pp. 523-534.

8. Oliver, W. L., J. M. Holt, R. H. Wand and J. V. Evans, Millstone Hill incoherent scatter observations of auroral convection over 60°<Λ<75°, 3. Average patterns versus Kp, *J. Geophys. Res.*, **88**, 1983, pp. 5505-5516.

9. Alcayde, D., G. Caudal and J. Fontanari, Convection electric fields and electrostatic potential over 61°<Λ<72° invariant latitude observed with the European incoherent scatter facility, *J. Geophys. Res.*, **91**, 1986, pp. 233-247.

10. Burke, W. J., D. A. Hardy, F. J. Rich, R. C. Sagalyn, B. Shuman, M. Smiddy, R. Vancour, P. J. L. Wildman, M. C. Kelley, M. Doyle, M. S. Gussenhoven and N. A. Saflekos, High latitude electrodynamics: Observations from S3-2, *Space Sci. Rev.*, **37**, 1984, pp. 161-200.

11. Killeen, T. L., R. A. Heelis, P. B. Hays, N. W. Spencer and W. B. Hanson, Neutral motions in the polar thermosphere for northward interplanetary magnetic field, *Geophys. Res. Lett.*, **12**, 1985, pp. 159-162.

12. Heelis, R. A., P. H. Reiff, J. D. Winningham and W. B. Hanson, Ionospheric convection signatures observed by DE 2 during northward interplanetary magnetic field, *J. Geophys. Res.*, **91**, 1986, pp. 5817-5830.

13. Burke, W. J., M. C. Kelley, R. C. Sagalyn, M. Smiddy and S. T. Lai, Polar cap electric field structures with a northward interplanetary magnetic field, *Geophys. Res. Lett.*, **6**, 1979, pp. 21-24.

14. Hoffman, R. A., M. Sugiura, N. C. Maynard, R. M. Candey, J. D. Craven and L. A. Frank, Electrodynamic patterns in the polar region during periods of extreme magnetic quiescence., sub. *J. Geophys. Res.*, 1988.

15. Iijima, T., and T. A. Potemra, Field-aligned currents in the dayside cusp observed by Triad, *J. Geophys. Res.*, **81**, 1976, pp. 5971-5979.

16. Zanetti, L. J., and T. A. Potemra, The relationship of Birkeland and ionospheric current systems to the interplanetary magnetic field, *Solar Wind-Magnetosphere Coupling*, edited Y. Kamide and J. A. Slavin, Tokyo, Terra Scientific Pub. Co., 1986, pp. 547-562.

17. Iijima, T., and T. A. Potemra, Large-scale characteristics of field-aligned currents associated with substorms, *J. Geophys Res.*, **83**, 1987, pp. 599-615.

18. Hoffman, R. A., M. Sugiura and N. C. Maynard, Current carriers for the field-aligned current system, *Adv. Space Res.*, **5**, 1985, pp. 109-126.

19. Hoffman, R. A., and M. Sugiura, Field-aligned currents as a function of substorm phase, IUGG/IAGA, XIX General Assembly, 1987, abstract, p. 589.

20. Fujii, R., R. A. Hoffman, and M. Sugiura, Spatial relationships between region 2 field-aligned currents and auroral particle precipitation in the evening sector, *EOS*, **68**, 1987, abstract, p. 1443.

21. Sugiura, Masahisa, A fundamental magnetosphere-ionosphere coupling mode involving field-aligned currents as deduced from DE 2 observations, *Geophys. Res. Lett.*, **11**, 1984, pp. 877-880.

22. Bythrow, P. F., T. A. Potemra, L. J. Zanetti, R. A. Erlander, D. A. Hardy, F. J. Rich, and M. H. Acuna, High latitude currents in the 0600 to 0900 MLT sector: observations from Viking and DMSP-F7, *Geophys. Res., Lett.*, **14**, 1987, pp. 423-426.

23. Fujii, R., R. A. Hoffman, M. Sugiura, J. D. Craven, L. A. Frank and N. C. Maynard, The field-aligned current system associated with the bulge-type auroral expansion, *EOS*, **69**, 1988, abstract, p. 445.

24. Reiff, P. H., H. L. Collin, J. D. Craven, J. L. Burch, J. D. Winningham, E. G. Shelley, L. A. Frank and M. A. Friedman, Determination of auroral electrostatic potentials using high- and low-altitude particle distributions, sub. *J. Geophys. Res.*, Jan. 1988.

25. Weimer, D. R., C. K. Goertz, D. A. Gurnett, N. C. Maynard and J. L. Burch, Auroral zone electric fields from DE 1 and DF 2 at magnetic conjunctions, *J. Geophys. Res.*, **90**, 1985, pp. 7479-7494.

26. Aggson, T. L., Probe measurements of electric fields in space, *Atmospheric Emissions*, edited by B. M. McCormac and A. Omholt, New York, Van Nostrand Reinhold, 1969, p. 305.

27. Maynard, N. C., A. Bahnsen, P. Christophersen, A. Egeland and R. Lundin, An example of anticorrelation of auroral particles and electric fields, *J. Geophys. Res.*, **78**, 1973, pp. 3976-3980.

28. Maynard, N. C., D. S. Evans, B. Maehlum and A. Egeland, Auroral vector electric field and particle comparisons: 1. Premidnight convection topology, *J. Geophys. Res.*, **82**, 1977, pp. 2227-2234.

29. Lyons, L. R., Generation of large-scale regions of auroral currents, electric potentials and precipitation by the divergence of the convection electric field, *J. Geophys. Res.*, **85**, 1980, pp. 17-24.

30. Lyons, L. R., Discrete aurora as the direct results of an inferred high-altitude generating potential distribution, *J. Geophys. Res.*, **86**, 1981, pp. 1-8.

31. Basu, Sunanda, Santimay Basu, E. McKenzie, 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**., 1988, 115-136.

32. Hoffman, R. A., Properties of low energy particle impacts in the polar domain in the dawn and dayside hours, in *Magnetosphere-Ionosphere Interactions*, edited by K. Folkestad, Oslo, Universitetsforlaget, 1972, pp. 117-138.

## ACKNOWLEDGEMENTS

The excellent manuscripts of the articles reviewing the scientific results during the first five years of operations of the Dynamics Explorer program, which were provided by L. A. Frank, J. D. Craven, J. L. Burch and R. A. Heelis, and which will appear in Reviews of Geophysics, were most timely and useful in the preparation of this review. J. D. Craven specially generated Figure 1 for this paper. R. A. Heelis and W. B. Hanson kindly furnished the unpublished ion drift data in Figure 7. R. M. Candey provided invaluable computer graphics assistance. The use of several figures from a manuscript in preparation by R. Fujii is gratefully acknowledged, as well as discussions with him. The unprecedented level of collaborations between members of the Dynamics Explorer Science Team has been a distinguishing characteristic of the program and the principal reason for the successes of the program.

## DISCUSSION

**T.Croft**

Have you looked at the South Pole? If so, are there any qualitative differences?

**Author's Reply**

Yes, I have looked at the South Pole. There are no significant differences from the standpoint of physics. Differences exist because the magnetic-to-geographic pole separation is greater there, and because the North-or-South $B$ field relationship is reversed.

# EMPIRICAL MODELS OF CONVECTION ELECTRIC FIELDS AND ELECTROSTATIC POTENTIAL AT HIGH LATITUDE FROM EISCAT OBSERVATIONS: PRELIMINARY RESULTS

by

Catherine Senior, Dominique Fontaine and Gérard Caudal
Centre de Recherches en Physique de l'Environnement
4, Avenue de Neptune
94107 Saint-Maur-des-Fossés
France

## ABSTRACT

Convection electric fields are observed with the EISCAT European Incoherent Scatter Common Program CP-3, which probes the $F_2$ region of the auroral ionosphere between 61° and 70° invariant latitude for 24-hour runs. Average convection patterns have been calculated from ion drift data gathered in several tens of these 24-hour experiments conducted between January 1982 and June 1987, for various levels of magnetospheric activity and orientation of the interplanetary magnetic field. The averaging technique is presented and the physical meaning of our criteria for selecting the data samples are discussed.

Electrostatic potential distributions are computed from these statistical electric field models, using the method developed by Alcaydé et al. (JGR, 91, 233, 1986). These electrostatic potential patterns are compared to earlier models obtained from incoherent scatter radar or satellite data.

Our results are consistent with the well-known two-cell convection pattern with antisunward flow over the polar cap, and sunward flow in the auroral zone. The characteristics and variability of these two cells are discussed in the light of indices of magnetospheric activity and orientation of the interplanetary magnetic field. Particular emphasis is placed on the shape and position of the two cells, and on the intensity of the dawn to dusk potential drop applied to the ionospheric auroral zone probed by the radar.

## DISCUSSION

**G.Rostoker**

I have one problem with treating the data the way you have, particularly in the vicinity of 2200-2400 universal time when you are taking a lot of data and putting it in bins. The electrojets and the convection zones move back and forth with varying levels of magnetic activity, so that if you were to take an instantaneous picture I think the probability is very high you would have much larger electric field vectors in that critical zone, i.e., in the Harang discontinuity just before midnight. The result is that the vectors average out to zero in the binning process, and you get very small vectors in a zone which I am sure should have very large vectors in it. I suspect the reason for this result is the movement of the zone.

**Author's Reply**

I think you are correct.

# LOW FREQUENCY ELECTROSTATIC WAVES OBSERVED IN THE VICINITY OF AN AURORAL ARC

by

K. Rinnert

Max-Planck-Institut für Aeronomie
D-3411 Katlenburg-Lindau
Fed. Rep. Germany

## ABSTRACT

Measurements of the electric field vector from d.c. to about 150 Hz were performed on board the CAESAR-II payload. The payload flew over a faint localized auroral arc associated with an inverted-V precipitation event. Equatorward of this auroral disturbance intense electrostatic low frequency signals were observed with fundamental frequency below the ion cyclotron frequency. These signals had properties of wave fields of ion-beam-driven electrostatic ion cyclotron waves (EIC) described by Yamada et al. [1]. The spatially or temporarily limited occurrence of these signals may give an indication of the distribution of ion beams or field-aligned currents.

## INTRODUCTION

To study auroral arc physics the CAESAR-program (Coordinated Auroral Experiment using Scatter And Rocket Investigations) was performed, which combined high resolution in-situ measurements using rocket-borne instrumentation and ground-based scatter observations [2,3] The CAESAR rocket was launched on January 30, 1985 at 19:30 UT from the Andoya rocket range under slightly disturbed conditions over a faint and slowly varying auroral arc. The payload reached its apogee of 703 km just above the arc, i.e. the trajectory crossed the magnetic field lines connected with the arc region near its apogee. The overall activity was low (Kp = 3-, Ap = 17) but the electron spectrometer data from the payload clearly showed increased particle fluxes for electrons with a typical inverted-V type of energy distribution between 410 s and 600 s flight time [3].

The electric field measuring instrument (a floating double probe system utilizing two crossed boom pairs) sampled the three-dimensional electric field every 3.2 ms and provided d.c. electric field data and a.c. electric field fluctuations with frequencies up to about 150 Hz. The resolution was about 1 mV $m^{-1}$. More details on this instrument are given in Ref. 2-4.

Synchronous with the CAESAR flight the EISCAT-system observed the plasma drift in an altitude of 140 km at seven locations of a south-north scan. From such drift measurements electric fields can be deduced. Assuming that electron-neutral collisions and neutral winds are negligible the electric field is derived from Eq. (1).

$$\overline{E} = \frac{\overline{v} \times \overline{B}}{B^2} \tag{1}$$

($\overline{v}$ is the plasma drift velocity, $4\overline{B}$ is the Earth magnetic field.)

Figure 1 shows the plasma drifts related to the 140 km altitude level in geographic coordinates. The squares labeled A - G indicate the locations where EISCAT measured the plasma drift vector. The dense sequence of vector arrows presents drift velocities calculated from CAESAR electric fields assuming pure $\overline{E} \times \overline{B}$ drift. The electric fields measured along the trajectory are mapped down the Earth magnetic field lines to the altitude of 140 km. (The foot points of these arrows represent the projected trajectory of the payload.) The full dots are time marks every 100 s flight time. South of the precipitation region (about 71.8°) both data sets are in excellent agreement and indicate a rather uniform strong westward drift corresponding to a northward directed electric field. At the southern boundary there is a well defined clockwise vortex (between the EISCAT positions E and F and the projected trajectory at about 450 s flight time), and at the northern boundary a counter clockwise vortex is indicated by the CAESAR data.

With respect to the Earth magnetic field the measurements on board the payload cover about two degrees of latitude south of the precipitation region during upleg. The downleg of the trajectory is close to parallel to the Earth magnetic field. During the rocket flight which lasted 14 minutes the electric field configuration was stable.

## MEASUREMENTS

South of the auroral arc in a region of primarily northward directed d.c. electric field with magnitude of up to 55 $mVm^{-1}$ low frequency electrostatic fluctuation were observed. Figure 2 represents in the upper panel a plot of the d.c. electric field component perpendicular to $\overline{B}$ versus flight time or altitude. This is the data used to calculate the plasma drift vectors plotted in Figure 1. The bottom panel shows the spectral distribution of the electric field component parallel to $\overline{B}$ of the a.c. electric field. Each trace represents rms values averaged over 10 Hz wide bands and one spin period. The occurrence is impulsive and the intensity is concentrated in the frequency band below 30 Hz. There is a distinct anticorrelation between the amplitudes of these a.c. electric fields and the flux of low energy (< 100 eV) precipitating electrons measured onboard the payload [3]. The a.c. magnetometer on board the CAESAR payload did not receive corresponding magnetic fluctuations, thus these signals are pure electrostatic. The a.c. electric field components perpendicular to $\overline{B}$ ($E_\perp$) are less intense than the components parallel to $\overline{B}$ ($E_\parallel$), however, show similar structure in time and frequency.

ure 3 gives an imagination of the variability of these signals. Each curve represents the squared Fourier transform of the electric field component parallel $\overline{B}$ (as in Fig. 2) of one full spin period and is an estimate of the spectral power density. The results for consecutive spin periods are plotted versus time for the active time interval from 200 s to 450 s flight time. The activity varies as already seen in Figure 2 with a maximum between 300 s and 350 s flight time and a minimum around 280 s flight time. The signals are short living, i.e. they last shorter than one spin period of 0.65 s, and individual spectra vary from one spin period to the other, even during the period of maximum activity. After 380 s flight time there are some big but isolated events.

From four periods of different activity marked by the arrows in Figure 2, bottom panel, estimates of power density spectra for all three components with respect to the Earth's magnetic field are presented in Figure 4. Each spectrum is an average over a 20 s flight time interval defined by the numbers at each trace. The signals or waves are predominantly polarized parallel to $\overline{B}$. Assuming that these are electrostatic waves they propagate also primarily parallel to $\overline{B}$. The spectral power density peaks in the 20 - 30 Hz band in all of these cases and during the more active periods there is evidence of harmonics. The components perpendicular to $\overline{B}$ show roughly the same frequency dependence, however, they are a factor of three or more below the field component parallel to $\overline{B}$.

Figure 5 demonstrates some more characteristics of these waves. Plotted is the a.c. signal of consecutive spin periods from an active period and a quiet period. Time runs from bottom to top and to the right. Each of the vertical traces covers one complete spin period of 0.65 s of the payload. The dotted traces represent the $E_{\parallel}$ component, the dashes represent magnitude and direction fo $E_{\perp}$ in a compass card type of presentation (north is up and east is to the right).

During the more quiet period (421 - 426 s flight time) the wave forms occur isolated and randomly distributed in time indicating that the sources are either temporal or local. It is quiet between the pulses. There is also no correlation with spin phase as one would expect to exist if the pulses were generated by some interaction of the payload with the ambient plasma.

During the more activ period (350 - 355 s flight time) there are also prominent wave forms of similar magnitude and duration. They are more irregular, probably because of considerable contribution of harmonics. There are no quiet periods between the big signals, but a kind of background noise with higher fundamental frequency of about twice that of the prominent wave forms.

The presentation of Figure 6 is an attempt of a three dimensional display of the electric field vector as seen from the rocket payload. Each trace covers one spin period and the signal is plotted for 9 spin periods from the more quiet region at the end just south of the arc (the first 5 traces are the same as in Figure 5 for the more quiet period). It clearly demonstrates the localized occurrence of these fields embedded in a low intensity background. This background obviously is not just noise but shows significant coherency.

## DISCUSSION

In the polar ionosphere a host of plasma instabilities can be generated under the influence of current systems or plasma density gradients and electric fields. Most of these observations are made from polar orbiting satellites from altitudes above about 1000 km, and often these fluctuations are associated with auroral arcs, with precipitating particles or large plasma density gradients close to arcs [5-8].

In the case under discussion, there was only a week local disturbance with a faint and slowly varying arc, and the observed electrostatic fluctuations were observed equatorward of that arc. Equatorward of discrete arcs where often diffuse aurora is observed caused by low flux of precipitating particles. Above the arc, at the auroral fieldlines, where often current driven instabilities are observed, no electric field fluctuations were detected in the frequency range covered by the instrument (see Fig. 1).

Kindel and Kennel [9] identified electrostatic ion cyclotron (EIC) wavs as the mode of lowest threshold of current driven instabilities. To destabilize EIC waves a critical electron drift velocity $v_c$ of the order of the thermal velocity with respect to the ions must be exceeded. For the altitude under discussion (300 km - 700 km) $O^+$ is the dominant ion, and electron temperature $T_e$ and ion temperature $T_i$ are of the same order of magnitude. For such conditions Kindel and Kennel [9] deduced the real part of the wave frequency $\omega_r$ to be slightly above the gyrofrequency $\Omega_i$ ($\Omega_i < \omega_r < 1.2\ \Omega_i$), and the wave vector to be primarily perpendicular to the magnetic field ($k_{\perp} \approx 10\ k_{\parallel}$).

The gyrofrequency of $O^+$ in the region under discussion is about 38 Hz. The maximum growth rate of $O^+$ - EIC waves would be about 45 Hz and the waves should propagate nearly perpendicular to $\overline{B}$. Of course, the received signals are Doppler shifted because of the payload's motion, however, this effect is small under these specific conditions and cannot explain the discrepancy. Furthermore the Doppler effect should cause a widening of the averaged spectra rather than a shift.

Such $H^+$ - $EIC$ waves are often observed by satellite-borne instruments in greater altitudes where $H^+$ dominates, but not much data is available from below 1000 km whre $O^+$ is the dominant ion [7,10,11].

Waves destabilized by the gradient drift instability also propagate perpendicular to both $\overline{B}$ and the gradient. In our case the wave vector is primarily parallel to $\overline{B}$ and the waves seem to be localized, thus these two types of instabilities most important in association with auroral phenomena cannot adequately explain the observations.

Some features of these signals as the main field direction, their typical magnitude of $\sim 10$ mV m$^{-1}$ and their temporal or spatial occurrence are typical for solitary waves. However, the duration of solitary waves is much shorter, a few milliseconds rather than the 60 - 80 ms of the wave forms under discussion, and so far they have been observed only at much greater heights above $\sim 6000$ km. There they are a common phenomenon of the auroral zone high altitudes and are well correlated with upward directed substantial field aligned currents and ion beams. As seen in Figure 1, the wave activity is observed south of main precipitation region and not at the auroral field lines.

Electrostatic ion cyclotron instabilities may also be destabilized by ion beams. This mechanism has been investigated in the laboratory [1,12-14]. Yamada et al. [1] considered a Maxwellian plasma in equilibrium with a confining magnetic field $\overline{B}$. An ion beam (with the beam ions the same as the target ions) was injected parallel to $\overline{B}$. If the beam velocity $u_b$ is larger than about 3 times the ion thermal velocity $v_i$ a plasma instability within the beam is destabilized. The wave phase velocity of this cyclotron-cyclotron mode is about half the beam velocity. Yamada et al. have calculated the growth rates of this instability for various ratios $u_b/v_i$. The instability starts with $\omega$ about half the ion gyrofrequency $\Omega_i$. With increasing beam velocity the frequency of maximum growth rate approaches $\Omega_i$ and harmonics occur for $u_b \geq 5 \cdot v_i$. The energy of the beam ions in our case would still be less than 1 eV. The plotted curves of the calculated growth rates versus wave frequency resemble very much the power spectra of the observed signals under discussion as shown in Figure 4.

With $O^+$ the dominant ion in the upper ionosphere under discussion and $T_e$ and $T_i$ of about 1500 K as measured on board the payload the ion thermal velocity $v_i$ is about 1200 ms$^{-1}$. Assuming a beam velocity $u_b$ of $4 \cdot v_i$ or about 4800 ms$^{-1}$ the destabilized cyclotron-cyclotron mode would have a wave phase velocity of about 2400 ms$^{-1}$, a fundamental frequency of about $\omega \sim 0.6 \cdot \Omega_i = 143$ s$^{-1}$ or 23 Hz (frequency of maximum growth rate) and a parallel wavelength of about 100 m. The fact that the resonant waves are confined to the interior of the beam is also in accordance with the spatial or temporal occurrence of the signal when the rocket payload crossed such a beam. The observed wave forms around 420 s flight time are most instructive because these are really isolated events (Fig. 6).

With this type of instability the observed wave forms can easily be described. Given an ion beam of ionospheric ions moving upward parallel $\overline{B}$ with the cyclotron-cyclotron mode already destabilized and assuming a frame of reference moving with the wave phase velocity, the wave appears as a stationary electric field configuration or space charge distribution. Such a situation is sketched in Figure 7. The cylinder at the left represents the ion beam with a very simple charge modulation. In this frame of reference the payload moves along the dotted line and samples the electric field as displayed at the right. A real example of an observed wave form from the more quiet period shown in Figure 5 was selected and the beam modulation was drawn to fit that wave form. The individual three-dimensional wave forms depend on the geometry of the trajectory with respect to the beam centre and the wave phase. Each observed isolated waveform corresponds to a close flyby to the centre of such a narrow ion beam. The duration of the waveforms is typically 60 - 80 ms. During this time the payload moves about 40 m perpendicular to $\overline{B}$, thus 40 m is the characteristic beam width ($v_\perp \cdot 80$ ms $= 540$ ms$^{-1} \cdot 0.08$ s $\sim 43$ m). This is extremely narrow and is only about 8 times the gyration radius of the $O^+$ ions.

In the more active region as displayed in Figure 5 the events are of the same magnitude as in the quiet region, may be more often, and are characterized by harmonics (see Fig. 4). In the proposed scenario this means the ion beams are found closer and their velocity $u_b$ must be higher. With increasing $u_b$ the frequency of maximum growth rate approached the cyclotron frequency of the ambient plasma and the beam instability is heavily damped by cyclotron damping. By this mechanism energy is transferred from the beam instability to the ambient plasma and ordinary $O^+$ - EIC waves are generated and are observed as the background fluctuation.

As mentioned above there is an anticorrelation of the electrostatic wave activity with the low energetic particle flux, i.e. there is a maximum in the flux of electrons of $< 300$ eV between 260 s and 310 s flight time whereas there is a minimum in the wave activity. Thus, there is evidence of a relation between particle precipitation and the generation of this electrostatic plasma instability. Furthermore in the entire region with wave activity there was a substantial perpendicular electric field. The nonuniform flux of precipitating energetic electrons causes gradients in the electron density of the lower ionosphere and, thus gradients in the Pedersen and Hall conductivities. Such local conductivity gradients together with the high perpendicular electric field ($\approx 50$ mV/m$^{-1}$) are necessarily related with corresponding field-aligned currents (see Eq. 2).

$$j_\parallel = \Sigma_p div \overline{E}_\perp + \overline{E}_\perp \cdot grad \overline{\Sigma}_p + \overline{B} \times \overline{E}_\perp grad \Sigma_H \qquad (2)$$

$\Sigma_p$ and $\Sigma_H$ is the height-integrated Pedersen and Hall conductivity, respectively.

## SUMMARY

Electrostatic burst-like wave fields have been observed south of a faint auroral arc in altitudes above $\sim 400$ km with fundamental frequency of about half the gyrofrequency of $O^+$ the dominant ion at that height in the ionosphere. These electric field fluctuations had properties of ion-beam-driven EIC waves described by Yamada and colleagues [1]. If this type of plasma instability is the proper explanation for the observations it would imply that very narrow and localized field aligned beams of upward moving ionospheric ions exist. Such beam instabilities may develop into solitary waves and double layers in greater heights. The combination of large $E_\perp$ and localized gradients in the ionospheric conductivity due to varying precipitation, may explain the development of field-aligned currents carried by ionospheric ions, however, the deduced narrowness of the beams still remains to be suspicious.

## REFERENCES


1 Yamada M., S. Seiler, H.W. Hendel and H. Ikezi, Electrostatic ion cyclotron instabilities dri.en by parallel ion beam injection, *Phys. Fluids* **20**, 450, 1977

2 Wilhelm K., Campaign Handbook, Vol. I - CAESAR Investigations, Scientific Aspects and Experiment Description, MPAE-W-47-83-21, Max-Planck-Institut für Aeronomie, Katlenburg-Lindau, FRG, 1983.

3 Wilhelm K., K. Rinnert, K. Schlegel, H. Kohl, N. Klöcker, H. Lühr, W. Oelschlägel, G. Dehmel, M.P. Gough, B. Holback and K.-I. Oyama, Co-ordinated Auroral Experiments Using Scatter and Rocket Investigatius, Final Report on the Scientific Aspects, MPAE-W-47-87-13, Max-Planck-Institut für Aeronomie, Katlenburg-Lindau, FRG, 1987.

4 Rinnert K., Instrument for Measuring Electric Fields in the Ionosphere (Instrument CL13 of CAESAR-Rocket Payload), MPAE-W-40-84-05, Max-Planck-Institut für Aeronomie, Katlenburg-Lindau, FRG, 1984.

5 Bering E.A., Apparent electrostatic ion cyclotron waves in the diffuse aurora, Geophys. Res. Lett. **10**, 647, 1983.

6 Gurnett D.A. and L.A. Frank, A region of intense plasma wave turbulence in auroral field lines, J. Geophys. Res. **82**, 1031, 1977.

7 Kintner P.M., M.C. Kelley, R.D. Sharp, A.G. Ghielmetti, M. Temerin, C. Cattell, P.F. Mizera and J.F. Fennel, Simultaneous obse.vations of energetic keV upstreaming ions and electrostatic hydrogen cyclotron waves, J. Geophys. Res. **84**, 7201, 1979.

8 Maggs J.E. and W. Lotko, Altitude Dependent Model of the Auroral Beam and Beam-Generated Electrostatic Noise, J. Geophys. Res. **86**, 3429, 1981.

9 Kindel J.M. and C.F. Kennel, Topside Current Instabilities, J. Geophys. Res. **76**, 3055, 1971.

10 Gorney D.J., A. Clarke, D. Croley, Jr. J. Fennell, J. Uhmann and P. Mizera, The Distribution of Ion Beams and Conics Below 8000 km, J. Geophys. Res. **86**, 83, 1981.

11 Mizera P.F., D.J. Gorney and J.F. Fennell, Experimental Verification of an S-Shaped Potential Structure, J. Geophys. Res. **87**, 1535, 1982.

12 Michelsen P., H.L. Pécseli, J.J. Rasmussen and N. Sato, Unstable electrostatic ion cyclotron waves excited by an ion beam, Phys. Fluids **19**, 453, 1976.

13 Ohnuma T., T. Fujita and S. Adachi, Two stream ion-ion instability in an axially bounded ion beam plasma system, Phys. Fluids **21**, 1551, 1978.

14 Miura A., H. Akuda and M. A..... ur-Abdalla, Ion-beam-driven electrostatic ion cyclotron instabilities, Geophys. Res. Lett. **10**, 353, 1983.




Figure 1. Plasma drift velocities in 140 km altitude measured by EISCAT at the positions A–G, and those derived from CAESAR electric field data measured along the trajectory and mapped down along the Earth magnetic field.

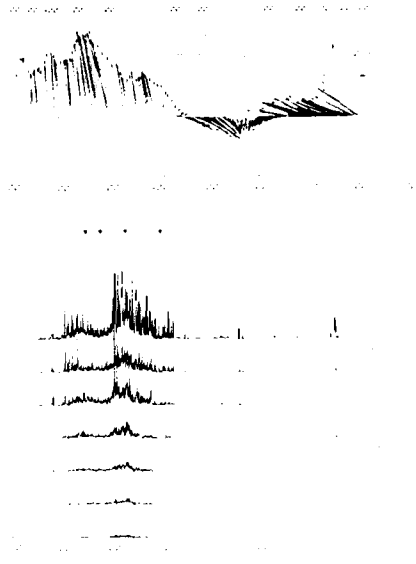


Figure 2. Electric fields measured along the CAESAR-II trajectory.
Top panel: d.c. electric field component perpendicular to the Earth's magnetic field $B$.
Bottom panel. Spectral distribution of the a.c. electric field component parallel to $\bar{B}$ averaged over 10 Hz wide bands as defined by the frequency intervals.



Figure 3. Estimates of spectral power density of the a.c. electric field component parallel $\overline{B}$ versus flight time. Each individual curve is the squared Fourier transform of the $E_{\parallel}$ signal for successive spin periods and represents an estimate of spectral power density (in arbitrary units). The spin modulated d.c. component has been subtracted.

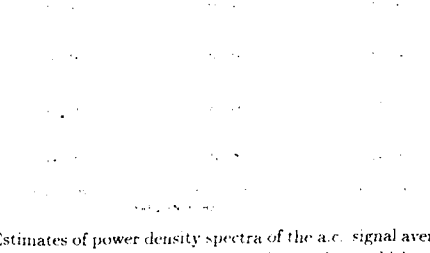


Figure 4. Estimates of power density spectra of the a.c. signal averaged over 20 second time intervals defined by the numbers, which are also marked by the arrows in Fig. 2, bottom panel. $E_{\parallel}$, $E_{\perp south}$ and $E_{\perp east}$ indicate the spectra of the electric field components parallel, southward directed and eastward directed with respect to the Earth magnetic field.

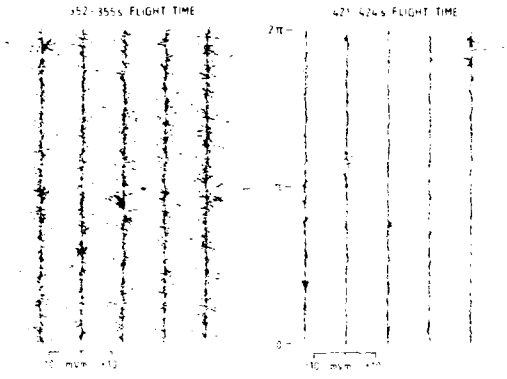



Figure 5. Sections of received a.c. signals from two flight time intervals with high and low wave activity. From each period five consecutive full spin periods are plotted. The dotted traces represent $E_{\parallel}$ and the dashes represent $E_{\perp}$ in magnitude and direction with respect to $\overline{B}$.



Figure 6. 3-dimensional plot of the electric field vector of the a.c. signals versus time. Each trace represents one spin period. UP means upward parallel to the Earth magnetic field $\overline{B}$, N and E mean north and east with respect to $\overline{B}$.

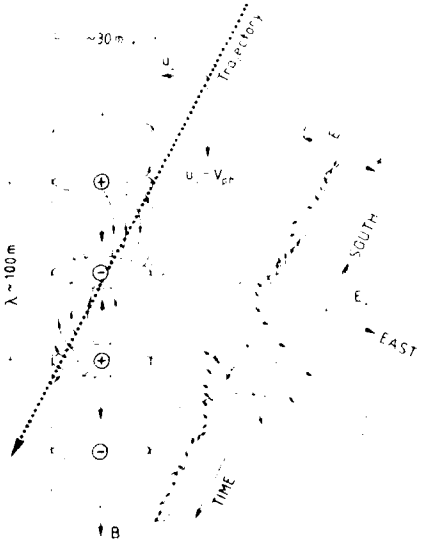



Figure 7. Sketch of an ion beam cylinder with the cyclotron-cyclotron mode destabilized. A simple charge density modulation is assumed within the cylinder and the arrows represent the associated electric field. In a frame of reference moving with the wave phase velocity of the instability the payload moves along the dotted line and samples the electric field as plotted at the right, which is a an example of an observed wave form.

## DISCUSSION

**G.Rostoker**

Do you have particle spectra (distribution functions) for this pass? I would like to know if the flux versus energy plots showed a drifting Maxwellian (representative of the bps) or power law (representative of the cps) signature. I ask this question because it seemed that at your highest frequencies there was no noise—thus I am not sure you are seeing a true inverted "V" structure with all the properties attributed to inverted "V's" by Frank et al.

**Author's Reply**

The particle spectra clearly show faintly narrow secondary peaks in the electron flux indicating acceleration along B typical for inverted-V events. The electric field measuring instrument is frequency band limited to $f \lesssim 1 < 0$ Hz. It does not detect higher frequency wave fields associated with inverted V's. The low frequency electrostatic wave fields described were detected south of the arc outside the inverted-V region. (There only soft precipitation $\sim 100$ eV was measured.)

# POLAR F LAYER "LACUNA" IRREGULARITIES AT DUMONT D'URVILLE

by

Paul Vila
EMI-RPE/RAB/CNET
38—40 rue du Général Leclerc
92131 Issy-les-Moulineaux
France

## ABSTRACT

Polar F lacunae, a loss of F-region echoes by the vertical-incidence ionosondes are re-investigated. The DUMONT D'URVILLE series including rapid-run sequences suggest the existence of three distinct classes of lacuna. Combining these morphological descriptions with some specific observations by backscatter radar, incoherent scatter meridian scans, angle-of-arrival ionosonde measurements and VIKING satellite, we suggest a dual mechanism for the propagation loss causing lacunae. This consists in a fine E—F sporadic screen structure sometimes identical to the OLESEN "SEC" phenomenon, and of F-region fronts about 300 km wide which tilt the scattered ionosonde waves away from the captation cone of the ionosonde receiving antenna. The scant physical measurements available for this study allow us to propose an unifying model for the front and screen formation, which needs checking up with more co-ordinated data.

## RESUME

On redéfinit les lacunes polaires, qui s'apparentent à un défaut de propagation des ionosondes verticales, fréquent aux heures de jour à l'intérieur de la calotte polaire d'été. Une approche empirique des traces d'ionogrammes indique deux mécanismes complémentaires de cette perte d'échos, liés à deux types de structures distincts:

(A) La structure fine d'écran dispersif — considérée à l'exclusion de toute autre par le groupe Danois — liée à une mince couche d'instabilité double-faisceau vers 110 km, ou "SEC". Ce modèle ne permet pas d'expliquer les lacunes F2, les plus fréquentes aux latitudes polaires, et on suggère que l'effet d'écran peut être attribué aux irrégularités E-sporadiques produites par gradient dérive entre 115 et 155 km dans la vallée d'ionisation entre les couches E et F1.

(B) Le phénomène de réflexion oblique in situ sur des fronts d'échelle 200 à 500 km horizontale, qui déflectent le signal hors de la zone de captation de l'antenne réceptrice, proposé par Michel Sylvain.

Les parts respectives de ces deux phénomènes varient selon les classes de lacunes, dont on indique les principales caractéristiques d'occurrence, de durée de vie et de traces diffuses à DUMONT d'URVILLE (Antarctique). Quelques mesures d'incidence oblique révèlent l'existence de la convexité (B) Grace à quelques données synoptiques du satellite VIKING, on relie les hypothèses précédentes aux schémas théoriques sur l'entretien des arcs et des fronts de plasma de la calotte polaire calme. Il est suggéré que la double structure de fronts F2 et d'écran E-F1 peut être la conséquence d'une même précipitation déposant des électrons de faible énergie à des altitudes basses de la couche F2, dont la structure fine d'instabilités de gradient x dérive se projette par diffusion dans la vallée E-F1.

## INTRODUCTION

### Definition

F lacuna is a *total* disappearance of the vertical incidence ionogram echoes from a *whole F layer*, be it the F1, the F2 layer or the entire F region — in which case we call it F3 lacuna. It is exclusively a summer daytime phenomenon of the inner Polar Cap ionosphere.

### "Slant E Conditions"

The above definition is disputed by the Danish Group at Copenhagen who first documented a particular case of F1 lacunation (Olesen and Rybner 1958 [1]) named Slant E Condition: indeed during active periods at auroral polar cap latitudes the Slant Sporadic E trace is often — but not always — accompanied by a "damping" of the F1 layer echoes, however generally incomplete. The 1958 paper interpreted brilliantly this damping as diffractive loss by a screen of type I instability irregularities at lower E layer level. But this interpretation cannot stand for all F1 lacunae, and becomes very difficult for the F2 types of lacuna.

The present paper is an attempt at clarifying the controversy, and towards a solution of the lacuna riddle, which will help us make full use of these phenomena for a better understanding of Polar Cap dynamics and electrodynamics. In our first section, we analyse lacuna occurrence from low-Sun-spot activity periods at Dumont d'Urville (67° S., 140° E., 80.5° Invariant Latitude): we outline the hourly occurrence and describe in more detail a series of rapid-run sequences obtained in 1975 (M.C.Lecomte, unpublished), completing the high Sunspot activity picture (Sylvain et al, 1978 [2]).

Our second section reviews recent results about the SEC-F1 lacuna geometry, from the Halley Bay advanced ionospheric sounder (A.Roger and M.Pinnock [3]). We compare their interpretation with that from experiments in

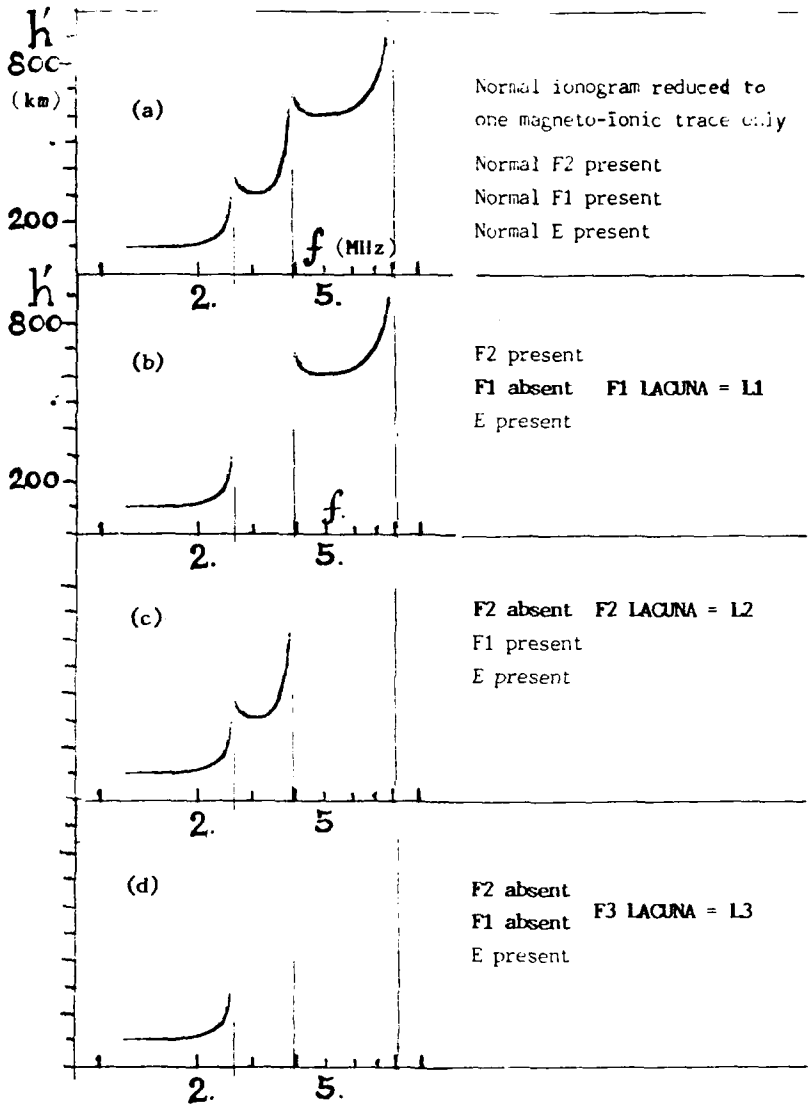



Figure 1 Simplified ionogram and the three lacuna cases

Greenland combining the Danish ionosondes and HF backscatter radar with meridian scans from the Sondrestrømfjord incoherent scatter facility (Olesen et al., [4], Kelly et al. [5]).

In a third section we show a few Viking satellite U.V. images of auroral and polar arcs. These pictures, obtained while lacuna conditions existed on the vertical ionosonde at Qanaq (Thule, 77.5°N, 69°W; ≅89° Invariant Latitude) reveal the great variability of auroral conditions associated with the lacuna phenomenon.

In the light of these data we discuss in a fourth section the theoretical contributions to the plasma structures producing lacunae. We appeal to the International community for the joint use of already existing experiments: the ionosonde network and the Qanak digisonde, the Sondrestrømfjord incoherent scatter sounder and possibly aeronomic polar satellites, which would definitely solve the polar F lacuna riddle.

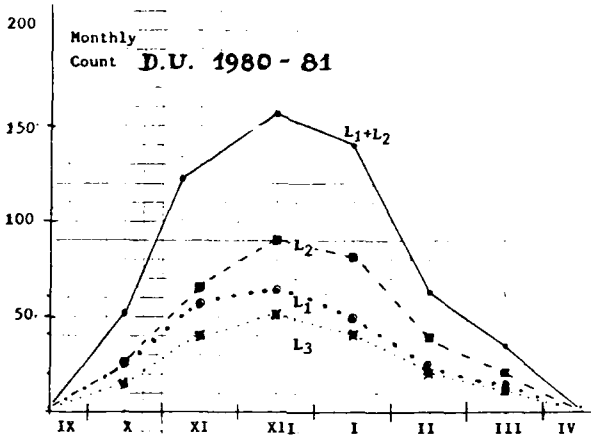



Figure 2 Seasonal variation

## I – DUMONT D'URVILLE LACUNA OCCURRENCE

### Hourly data

Figure 2 shows the month-by-month number occurrence for 1980–1981; we plot separate curves for the three types of days with only one class of lacunae, either F1, F2 or F3, and for the mixed-class type F1 + F2 days. They show identical trends with a maximum at summer solstice. Figure 3 shows the cumulated diurnal curve for all months together, peaking very near magnetic noon, with a slightly earlier rise in the F2 class; for first lacuna observed in the morning during mixed-class days is an L2, especially on quiet periods; this suggests a simpler set of conditions during these early local times when the F2 layer plasma is only driven by steady convection, and structured by low-energy electrons.

A statistical index has been derived which measures for each individual day the departure of the number of lacunae from monthly average. Its distribution in function of the relevant magnetic characters is revealing: F1 lacunae *by themselves* are limited to the *least magnetically active* and least lacuna crowded days. This is exactly the opposite during high Sunspot activity years when magnetosphere and ionosphere plasma contents are larger; we suggest that penetration at E-F1 levels by the particle fluxes requires higher energy, hence higher magnetic activity at Sunspot maximum. In the same occurrence change versus magnetic activity domain, F2 lacunae cover a wide range of average conditions, while F3 cases become more frequent with higher — non storm — magnetic activity. During strong magnetic disturbances, all ionogram traces disappear by D region absorption Black-Out, often preceded by strong sporadic E occultation due to KeV energy particle precipitation.

We have re-considered as critically as possible the basic work of Sylvain et al [2], and satisfied ourselves by frequent checks on the data about the frequent existence of F2 lacuna as free from "impeachment" G condition of the F2 layer peak $f_0F2 < f_0F1$.

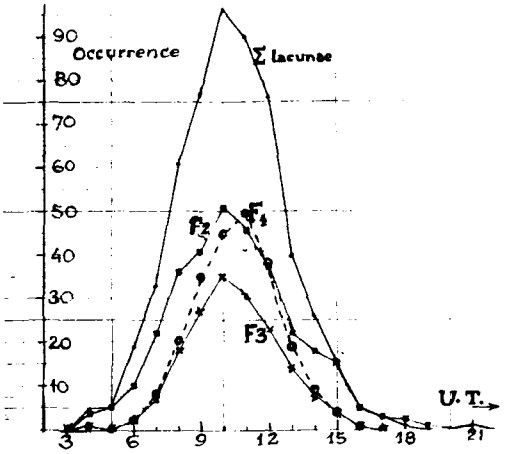



Figure 3 Cumulated diurnal distribution

A simple study of partial lacunarity versus true lacuna reveals that although lacuna F traces vanish, the plasma density overhead does not: it is slightly distorted in the F1 valley where the conditions of transition between collisions with neutrals and gyromagnetic field-aligned paths for the ions $\nu_{in} = \omega_i$ are reached. At F layer heights the plasma can also be modulated at hundreds of km scales, and transitions from underdense to overdense F2 layer peak affect F2 and F3 lacunae, as we will see later. The statistical comparison has been made for a typically crowded month, October 1974.

For this sample Figure 4a shows the percentage plot of all F region ionogram parameters across altitude space (hourly data) replaced by Y symbols:

– at 08. LT, 55% lacunarity cases affect the F2 layer
29% full F2 lacunae are present
16% complete F3 lacunae arise.

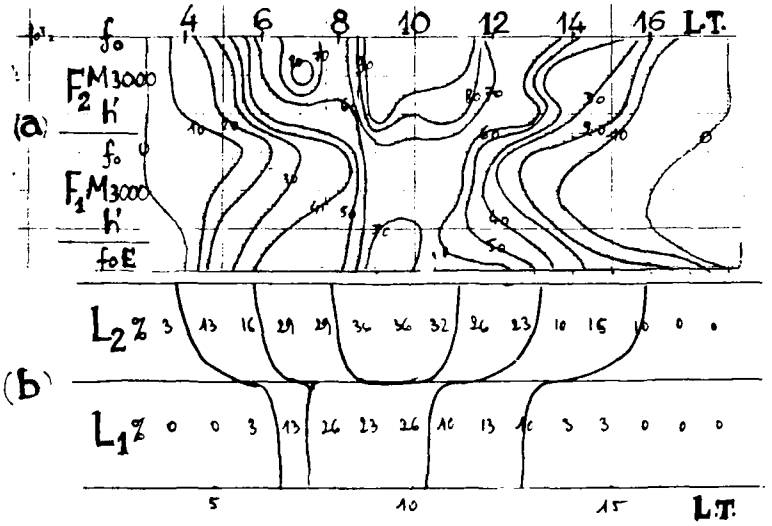



Figure 4

(a) Percentage field of lacunation (Y symbols = no echo observed) on the October 1974 hourly scalings at Dumont d'Urville interpolated in virtual height, versus Local Time, from the three parameters: $-f_o$ = peak frequency $M_{3000}$ = midlayer profile shape characteristic $-h'$ = minimum virtual height

(b) Percentages of full lacunae.

On Figure 4b the same plots for Lac F1 and Lac F2 are smoothed up for the two grades of phenomena:

(a) partial, or lacunarity
(b) full lacuna.

Both corolla-like patterns have similar shape, only with higher densities in (a); also (a)/(b) ratios, larger than 4 during post-sunrise, drop under 2 from about 07 to 14 LT. Owing to the monotonic electron density variation in E layer, this non-gradual change seems to reveal a specific state associated with lacunae, which are relatively more frequent than would be a simple prolongation of partial lacunarity. This state is seen earlier in LT and more drastically in F2 lacunae.

What is it? All the data suggest that F1 lacunae are accompanied by an underlying screen of irregularities: sometimes the 110 km SEC structure of Olesen et al, [4] more often a high-altitude sporadic E with possibly a faint trace when embedded inside the E-F1 valley of ionization. For F2 lacunae, the Syvain model of convex iso-ionic surfaces appears also necessary. The various grades of lacunarity would result from the geometrical conditions of the propagation. This semi-artefact character of lacunae would well correspond with the near verticality of the magnetic lines of force for the high-altitude screen, and with the polar cap plasma transport for the F2 layer tilts.

At this stage we must take into account the instrumental overall gain of the sounding system which modifies lacuna threshold. The differences between instruments are real, but modern ionosonde receivers are tuned to similar sensitivity in order to reveal the spreading against background noise without being too often saturated. In this respect the technical constraints at Dumont d'Urville and Thule are less stringent than those of near-auroral latitude stations like Godhaven or Narssarsuaq where multiple-scale spreading must impose fine gain tuning.

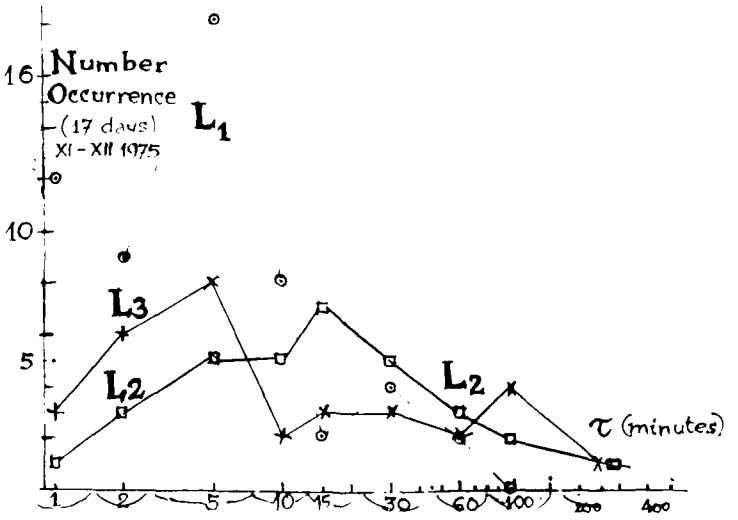



Figure 5

**Rapid-run sequences**

*Lacuna duration*

Figure 5 gives homemade spectra of the lifetimes of all lacuna events with one separate curve per class of lacuna. It shows much shorter durations than expected from quarter-hourly series. For F1 lacunae the main peak is on 5 minutes, with a secondary maximum on 30 minutes; even another peak at less than one minute is detected. The F2 lacuna distribution is smoother with a maximum around 17 minutes extending over 25 minutes. The L3 distribution shows a double population: one centred on the 5 minute L1 peak and a broad-band of long-lived events in the 1-to-3 hours domain.

Obviously the short-lived cases must be dominated by recombination processes near 120 km altitude, while the hour-long durations correspond to the cumulative effects of convective transport from the local noon area to the site. Less obvious is the mechanism maintaining F2 lacuna for nearly 20 minutes if an underlying screen at 110 km is the only responsible. We have to find maintenance processes for this screen in the L3 case anyway.

*Precise lacuna occurrence*

The two periods used, 8 days at the end of November and 7 days in December were quite different for the auroral conditions; this gives significance to simple averages through the sample. The results are given in Table I. They give a

weighted description for the lifetimes, and an idea of the change from 1975 to 1980, when corresponding daily averages numbers were only half as much.

With such short characteristic times, rapid sounding sequences prove necessary for catching up the generation process. The E-F1 desity gap traduced by ($f_oF1 - f_oE$) shows a consistent decrease of 5 to 20% immediately before lacunae, which lasts during F2 lacunae. The vertical true height extent of the valley increases at the same time; this can be explained if the valley, flattened in density and extended in height becomes liable to cross-field distortion.

**TABLE I Average lifetime and daily occurrence**

| | *L1* | *L2* | *L3* | |
|---|---|---|---|---|
| Duration *minutes* | 7.3 | 14.4 | 26.3 | |
| Number per day | 2.6 | 2.7 | 2.9 | TOTAL ≅8 |

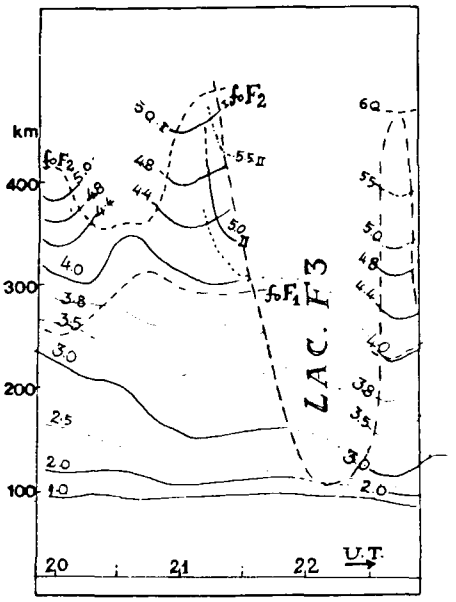



Figure 6

## II – NEW EXPERIMENTS

### The Advanced Ionospheric Sounder

Alan Rodger and M.Pinnock [3] evaluate angles of arrival with their array of dipoles at Halley Bay at various elevation angles. They sometimes observe slant sporadic E with F1 lacuna – when the auroral oval has expanded sufficiently far equatorwards. During the development of a lacunarity event on 26 September 1982 their residual echoes came from northern (equatorward) azimuths under gradually lower elevation angles. They concluded in favour of a combined mechanism 1°) the slant sporadic E irregularities scatter obliquely the upgoing waves; 2°) the tilted F1 reflector sends the sounding waves away from the site. This appears as a one-case confirmation of the Sylvain hypothesis about ray defocusing by F region patches during lacuna.

These results have been disputed by J.K.Olesen [6] who prefers to restrict all lacuna processes to Slant E Conditions alone and all slant E to his new model with electric field structures generated by heat depositions at low E region level. He favours quasi-vertical propagation with diffractive "damping" through the slant sporadic E array of irregularities. However his last study (J.D.Kelly et al. [5]) shows sporadic E reinforced patches detected by the incoherent scatter scans under the large-scale F region billows of underdense to overdense plasma travelling from polar to equatorial azimuths over Sondrefjord.

Although much too scanty and related to partial lacunation, these precious data point out to the same dual picture of the E-F1 dispersive screen, with a "blob" or patch convex F layer structure tilting the propagation, as intuitively suggested by our Dumont d'Urville series. The 26 September event at Halley Bay was probably an F3 lacunation. Conversely, J.K. Olesen may be right for some F1 lacunae about non-oblique propagation.

At Dumont d'Urville we have been able to approach the abrupt formation of an F3 lacuna on 13 February 1982. Figure 6 is a true height-versus-time plot of equal plasma density curves across the event of 22.UT ± 20 minutes. Here the tilts preceding lacunæ were so clear that the oblique traces could be followed through the 200 km to 450 km range of heights during nearly 30 minutes, and the upper F1 layer trace was also semi-oblique.

The precise dynamics are still missing.

Lacuna probing needs Incoherent scatter radar scans to the extreme ranges Northward from Sondrefjord, and Digisonde simultaneous tracking fo the E-F1 structures and gradients around Qanaq.

## III – POLAR ARCS AND LOW ENERGY PARTICLES SYNOPTICS

We have fortunately received some Viking Ultra-Violet pictures covering the polar cap at times nearly coincident with lacuna observations on Thule ionograms (we owe much gratitude to Dr J.K.Olesen for having been able to peruse them in Copenhagen two years ago). We describe these Viking pictures and some data recorded together on board of the satellite

26 July 1986 is a steady active day (Ap = 20, kp has risen from the day before to an average 3+) Thule magnetogram shows regular pc1 pulsations of 200 to 300 s. period and a few nT altitude, modulated by persistent larger oscillations of quasi-period 40 ± 20 minutes and 50 to 100 nT amplitude.

The Thule ionogram show a background of spurs and spread-F echoes indicating strong local ionisation gradients of minor amplitude, waving and waning from 08.30 to 20.00 LT. This background suddenly changes to give way to a lacuna sequence from shortly before 12.00 LT to 13.00 LT (= 14.56 to 16.10 UT).

Four Viking maps were selected, from positions of the satellite line-of-force closest from Thule for near-lacuna times, 13.45, 14.09, 17.55 and 18.19 UT. The stable feature on these maps is a thin dayside arc of continuous weak intensity, without cleft gap, located near the 75° invariant latitude circle. It remains shrouded on its North and South sides by a much wider and weaker diffuse luminosity; the Viking investigators at Calgary, Drs J.S.Murphree and R.D.Elphinstone [9], attribute this diffuse light to cloud contamination; however, additional arc branches come out of it on the 14.09 UT picture; later on, a series of thin arcs develop, "parallel to the noon oval and move toward the pole, fading within a couple of minutes". The Calgary authors mention that Dr G.Atkinson [10] has interpreted these short-lived features as originating from the plasma mantle boundary region during an "X line magnetospheric merging process near noon, from 13.35 to 13.41 and near 1400 and 1822 UT".

July 27 1986, a day of relatively decreasing activity (Ap = 16, average kp 4 in the morning down to 2 in the afternoon UT intervals) shows more substorm-like negative bays, 100 nT deep on the D and H components at Thule, less intense at Godhaven. The main lacuna event starts shortly before 14.30 UT and lasts until 21.00 UT.

The Viking pictures are too crowded for a clear distinction of any discrete features. However on the 16.09.57 map, the satellite was flowing overhead of Thule line-of force footprint, on the poleward edge of an inner band of East-West patches, the Viking electric field sensor enters a dynamic spiky area 20 minutes before lacuna; this long F2 lacuna series seems to be related to the stable band of sun-eastward patches which are probably sweeping the inner side of the afternoon convection cell. Intense electron fluxes of 100 to 700 eV energies arose at 16.10 UT, about 20 minutes before the lacuna. The Viking particles detectors recorded narrow ion cones with upgoing electrons, characteristic of parallel current circuits

By the end of the L2 event at Thule, the satellite approched Northern Greenland again between 20.17 and 20.40 UT However the nearer polar brightness was a faint distant patch. The low energy particle fluxes recorded equatorward of the bright auroral arc, in the 100 to 500 eV energy band were not surely present on the Thule line of force but the upward-accelerated electrons and the upgoing ions seem to characterize double-layer plasma beams. The earth magnetic field was becoming very weak, while the IBM By component started increasing eastwards. By this time it cannot be decided whether the F2 layer was experiencing a G situation, or if some tilted structure remained overhead. Lacuna sequences are much better understood on more frequent schedules; ground-based soundings with oblique direction-of-arrival would have helped clearing up the situation.

28 July 1986 is a quiet day (Ap = 9, kp average = 2.0) The satellite flies poleward at about 14.00 UT across a well-developed dayside arc on 75° magnetic latitude. The UV imager detects a large patch growing on either edge of this arc near the 11.00 LT meridian. Inside the polar cap the precipitative population flux densities are practically nil. Indeed the polar patch coincides with a typical polar plasma blob surging antisunward into the morning side of the cusp towards the polar area.The question remains of what maintains the F2 lacuna during the following three hours; for these hours, Sondre'fjord incoherent scatter meridian scans would probably tell how the blob source mechanism subsisted.

These few cases, an incomplete sample of the physics of F2 lacuna environment still leave many questions unanswered. However the precise Viking UV maps confirm in their variety the model derived from the previous morphological soundings: F2 lacunae are associated with clear discontinuities or arc-like plasma structures of the polar ionosphere, within a range of quite different geophysical configurations.

– On 26 July 1989, unstable Northward-traveling arcs.

— On 27 July 1986, an apparent meander of the afternoon convection cell border probably due to a resonant growth of substorm distortion across the oval.

— On 28 July 1986, an anti-sunward moving discontinuity identical with the quiet time blob of the plasma convection models.

The common topology in these various signatures is a locally stable gradient structure of the F region plasma.

## ACKNOWLEDGEMENTS

I wish to thank Ms Suzanne Cartron for inventing the polar lacunae on the first IGY ionogram series from Dumont d'Urville, Ms Marie-Christine Riouais-Lecomte for the thousands of minute-schedule ionograms she scaled, and Dr Michel Sylvain who passed me his rich files. Dr J.K.Olesen welcomed me at Copenhagen and let me see his July 1986 Thule ionograms, giving me the hope to see through our differences. Drs Murphree and Elphinstone processed and sent the Viking UV imager pictures just in time, and Dr Hervé de Feraudy helped me with them and with the accompanying data.

## REFERENCES

[1] Olesen J.K. and J.Rybner, AGARD-AG-34, 37-57, 1958.

[2] Sylvain M., J.J.Berthelier, J.Lavergnat and J.Vassal, Planet. Space Sci., 26(8), 785—799, 1978.

[3] Rodger A. and M.Pinnock, INAG Bul 48, August 1987.

[4] Olesen J.K. and P.Stauning, R.T.Tsunoda, Radio Sci., 127—140, 1987 and references to the controversy.

[5] Kelly J.D., R.T.Tsunoda, J.K.Olesen and P.Stauning, On the relationship of F region structure to HF Backscatter signatures attributed to the Polar Cusp, preprint communication.

[6] Olesen J.K, INAG Bull. 50, August 1988

[7] Heelis R.A., Vickrey J.F., and Walker N.B., J.Geophys. Res., 90, 437, 1985.

[8] Tsunoda R.T., High latitude F-region irregularities: a review and synthesis, Physica Scripta, 1987.

[9] Murphree J.S. and R.D.Elphinstone, private communication.

[10] Atkinson G., private communication.

# SEASONAL AND GEOMAGNETIC RESPONSE OF THE THERMOSPHERE AND IONOSPHERE.

D. REES and T.J. FULLER-ROWELL,

DEPARTMENT OF PHYSICS AND ASTRONOMY,
UNIVERSITY COLLEGE LONDON,
GOWER ST., LONDON WC1 E6BT.
UK

SUMMARY

The numerical and computational merging of the University College London Global Thermospheric Model and the Sheffield University Ionospheric Model has produced a self-consistent coupled thermospheric / ionospheric model, a valuable diagnostic tool for examining thermospheric / ionospheric interactions. The neutral thermospheric wind velocity, composition, density, and energy budget are computed, including their full interactions with the high-latitude ion drift and the evolution of the plasma density. A series of simulations has been performed at high solar activity, for a level of moderate geomagnetic activity, for each of the June and December solstices, and positive and negative values of the IMF-BY component. In the winter polar region, ion transport and the diurnal migration of the polar convection pattern into and out of sunlight play a major role in the plasma density structure at F-region altitudes. In the summer polar region, an increase in the proportion of molecular to atomic species, created by the global seasonal thermospheric circulation and augmented by the geomagnetic forcing, controls the plasma densities at all Universal Times. The increased destruction of F-region ions in the summer polar region reduces the mean level of ionization to similar mean levels seen in winter, despite the increased level of solar insolation. In the winter polar region at 300 km the dominant ion is $O^+$; in summer molecular and $O^+$ ions are of similar number densities. The summer ion temperature at 300 km exceeds the winter values by 500K, due to change in neutral temperature. In the lower thermosphere auroral oval the ion density is dominated by auroral precipitation in both seasons, resulting in only a small seasonal dependence in the height-integrated Joule heating rate and field-aligned currents (FAC). Within the polar cap, solar ionization generates a large seasonal variation of conductivity, producing a threefold increase in peak Joule heating rates, changing the balance of FAC. Neglect of neutral winds increases dusk sector Joule heating. Most of the neutral and electrodynamic parameters considered also have strong IMF-BY dependence.

## 1. INTRODUCTION.

This review will summarise recent advances in the development of coupled numerical models of the terrestrial thermosphere and ionosphere system. The behaviour of the high latitude regions will be the major topic, since that is where the influences of the magnetosphere are most strongly imprinted. The present status of the development of coupled models of the solar-terrestrial system will also be discussed, identifying the major outstanding problems of concept and theory.

At all times and at all geomagnetic activity levels, the magnetosphere imprints unmistakeable signatures upon the high-latitude thermosphere and ionosphere. The regions of such imprints recede poleward, away from areas of significant population, and thus historical observations, when the magnetosphere is in a 'quiescent' state. However, recent global observations from satellites, and new ground-based polar cap observatories have shown clearly that ionospheric structures, and thermospheric winds, temperature and composition are persistently disturbed in the vicinity of the auroral oval and within the geomagnetic polar cap. These disturbances are always co-located with signatures of energetic particle precipitation and convective electric fields, or else represent "fossils" of strong forcing during the previous 1 - 24 hour period.

A survey of the nature of the magnetospheric input and consequent thermospheric and ionospheric response, from quiet to very disturbed conditions, will thus be presented. The auroral oval generally expands equatorward and broadens as the level of geomagnetic activity increases. This expansion, broadening and an intensification of auroral precipitation is well shown in statistical surveys and analyses of the energetic electron precipitation /1,2/. These statistical surveys complement the impression obtained from individual observations. As shown by analyses of polar plasma convection /3-5/, the regions of strong magnetospheric convection electric fields imprinted on the polar ionosphere undergo a similar and closely related equatorward expansion and intensification as magnetospheric activity increases. From the point of view of numerical modelling, it is critically important that convection and precipitation boundaries match with reasonable fidelity, particularly if time-dependent simulations for variable geomagnetic activity are to be performed. Convection electric fields drive ionospheric plasma, within the auroral oval and polar cap, to velocities of the order of 1 km/sec. The ions impart momentum to the neutral gas via 'ion drag', at the same time losing a little of their net ($\underline{E} \times \underline{B}$) velocity, creating the dissipative Pedersen ionospheric current component, which causes Joule heating. If

there is to be any effective energy or momentum transfer from the solar wind, via the magnetopause and magnetosphere to the ionosphere and thermosphere, A/C and D/C components of the field-aligned current (FAC, or Birkeland current) are required. Such magnetospheric currents are associated with the dissipative Pedersen current within the auroral ionosphere (component parallel to the ionospheric electric field). Both the FAC and the Pedersen currents within the auroral ionosphere intensify sharply as geomagnetic activity increases. The efficiency of momentum transfer and the Joule heating both increase linearly with ionospheric plasma density. Since Joule heating almost always considerably exceeds direct particle precipitation, knowledge of the ionospheric plasma response to precipitation is particularly important. The polar regions which display the imprints of these important phenomena of magnetospheric origin also show a wide range of other disturbances. These disturbances produce a range of characteristic signatures in the charged and energetic particle populations, the AC and DC electric and magnetic fields, and the optical aurora. However, for the moment, we will concentrate on those phenomena which have the most direct connection to excitation of the thermosphere.

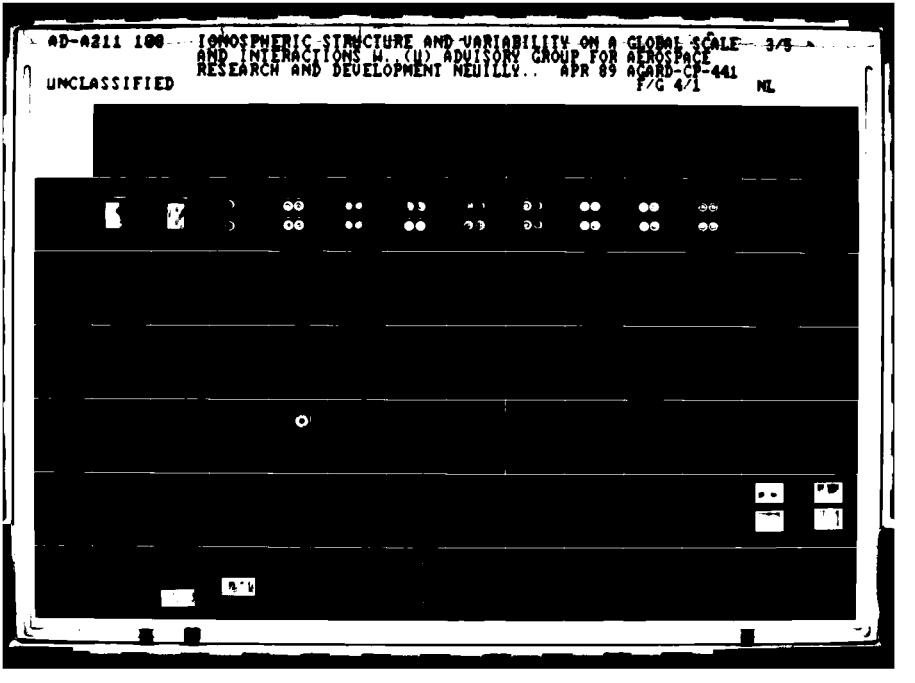
It is very important to recall that the classical ground-based signatures of intense auroral substorms: strong short negative magnetic excursions, brilliant aurorae and riometer observations of absorption events, are only a small part of the sequence of phenomena which cause strong ionospheric and thermospheric disturbances during geomagnetically disturbed periods. Indeed the signatures of the most important momentum and energy sources for the thermosphere are difficult to sense except by satellite or ground-based remote sensing. For these reasons, major thermospheric and ionospheric disturbances are often poorly related to classical indices such as Kp and AE, which are dominated by the magnetic effects of Hall currents during auroral substorms. Pedersen currents more accurately reflect the intensity of momentum and energy transport from the magnetosphere, but these, and the FAC system are poorly reflected in ground-based magnetic perturbations.

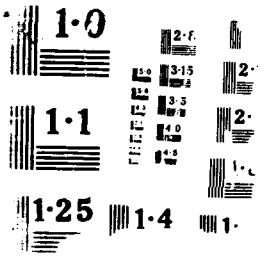
The signature of increasing geomagnetic activity is intensification of convective electric fields and precipitation, and an equatorward expansion of the regions affected (i.e. the auroral oval). These features also imply that FAC, Joule heating and ion-drag wind acceleration all increase. The temporal and spatial variability of all of the previously-mentioned terms, broadly speaking - 'geomagnetic forcing' - therefore increases, particularly at very high activity levels. Disturbances of the lower regions of the ionosphere, up to the normal E-region peak near 105-120 km, respond directly to 'auroral' inputs. However, strong E-region winds induced by ion-drag forcing generate a back-EMF opposing the convection electric field. At higher levels of the thermosphere and ionosphere, the situation becomes much more complex. Large scale advection and convection forced upon the thermosphere by geomagnetic heating causes the F-region neutral gas composition to change dramatically. There may be a very strong enhancement of molecular nitrogen density, and a corresponding depletion of atomic oxygen density. Coupled with dynamical effects on the F-region, resulting from strong induced horizontal winds, the ionospheric response to magnetospheric forcing becomes very complex and non-linear. Direct interpretation, without understanding the detailed neutral atmospheric response in the polar regions, is virtually impossible. Global-scale disturbances within the thermosphere follow initial high-latitude geomagnetic forcing. Propagating waves, and the consequences of gross wind-driven compositional changes have truly global consequences for the thermosphere, and force some very large, and long-lasting, disturbances of the ionosphere. During major geomagnetic disturbances, the decay of energetic particles from the ring current, probably mainly energetic ions /6/ appears to directly cause low-latitude disturbances of the thermosphere and ionosphere - the negative ionospheric storm. This process is, however, still rather difficult to include in a coherent and self-consistent model of the entire coupled Solar - Terrestrial system.

To understand the thermosphere, it is necessary to cope with the fragmentary nature of available empirical data. A major feature of the data on any parameter is that thermospheric variations are large. Many perturbations can be related to 'regular' diurnal or seasonal variations, to solar activity or geomagnetic activity factors, or to lower atmosphere tidal, gravity or planetary wave disturbances. However, the combination of sources, the unpredictability of several of the largest sources, the large magnitude of their induced perturbations, and the non-linear interaction between many of their effects cause the thermosphere to often appear quite chaotic. Thermospheric climatology is predictable, but thermospheric 'weather' can only be predicted by the likely magnitude of disturbances in certain regions, and within certain time periods. In particular, magnetospheric driving forces are complex, and the response of the stratified, always changing, thermospheric fluid is complex and non-linear. There are also a number of poorly-understood feedback mechanisms involving the plasma of the ionosphere, embedded within and originating from the thermospheric fluid. Numerical models provide a global framework incorporating the basic, well-understood, mechanisms and phenomena. Such models need to be global and time-dependent to be useful, since that is the nature of the thermosphere. Since we are not sure our understanding of all the thermospheric driving forces is complete, and our empirical description of the temporal and spatial dependence of even the major solar and magnetospheric driving forces is still quite inadequate, numerical models have to be used wisely. They provide a good means of predicting the mean structure, and of the qualitative and quantitative variations caused by seasonal changes and by solar activity variations. The general form of major large-scale thermospheric disturbances, the signatures of major magnetospheric disturbances, can be simulated quite well.

However, numerical models have to be used carefully in the prediction of thermospheric meteorology, particularly when small-scale and short-period disturbances are of interest. This is not particularly due to any problems of the models themselves, but simply reflects

AD-A211 100 IONOSPHERIC STRUCTURE AND VARIABILITY ON A GLOBAL SCALE 3/3
AND INTERACTIONS W.. (U) ADVISORY GROUP FOR AEROSPACE
RESEARCH AND DEVELOPMENT NEUILLY.. APR 89 AGARD-CP-441
UNCLASSIFIED F/G 4/1 NL




1·0
1·1
1·25
1·4

our lack of meteorological information on the variations of the magnetosphere, through which most solar wind inputs are imparted, directly or otherwise, to the thermosphere.

Numerical simulation of the thermosphere from first principles requires that the most important physical processes occurring within the thermosphere are properly treated /7-11/. It can be assumed that most of the energy and momentum sources driving the thermosphere are predetermined, and invariant to the response of the thermosphere. The thermosphere does not determine the nature of the solar UV and EUV inputs which provide important heat and ionisation sources. However, the thermosphere does react strongly to forcing. The major responses in wind, temperature and composition of the polar thermosphere to ion convection and heating within the auroral oval and polar cap are now well documented by ground-based and spaceborne observation /12-15/. Some of these thermospheric responses may change the nature or magnitude of the forcing itself. While this review will concentrate on external thermospheric forcing from the magnetosphere, it should be noted that significant effects, particularly in the lower thermosphere, occur from internal forcing from the lower atmosphere, as a result of the combination of tidal waves, gravity waves and planetary waves. The first can be handled numerically within a thermospheric model by introducing a 'flexible' lower boundary, where self-consistent changes of wind and temperature correspond to the amplitudes and phases of specific propagating tidal modes. The actual propagation of such tides through the lower and middle atmosphere is not simulated by such a procedure, however, the amplitudes and phases of several modes can be adjusted, by experiment, until the results within the lower thermosphere correspond to observed variations of tidal phases in temperature and wind (etc) as functions of altitude, season and latitude. It is found as would be expected, that the relatively large amplitudes of observed tidal winds in the lower thermosphere can be successfully simulated by introducing such propagating tides. This is not possible, if only the in-situ generated tides are considered. Additionally, large non-linear interactions occur between the individual propagating tidal modes themselves, and between them and the in-situ tides. While these studies still have to be considered as experimental (since the real tidal and meteorological behaviour of the middle atmosphere is still virtually unknown), these studies /16/ do provide a very useful insight into the complexities of the thermosphere.

The fundamental influences of the magnetosphere on the thermosphere come from magnetospheric plasma convection. This convection drives plasma into rapid motion within the polar cap and auroral oval. The structure and variability of polar convective fields were not well understood until the late 1960s and 1970s, with the advent of rocket and satellite measurements of ion drifts and electric fields /3,4/. These rapid plasma motions accelerate thermospheric winds via the process commonly known as ion drag /17/.

The resulting ion-neutral frictional drag causes direct heating of both ions and neutrals, commonly known as Joule heating /18,19/. Induced winds may increase or decrease (but generally decrease) the ion drag, and the resulting frictional heating. The induced winds (or more correctly, changed winds, since there is always a complex wind system in existence prior to a given geomagnetic disturbance) may induce a 'back-EMF', opposing the initial magnetospheric convective electric field. This wind system (subtly modified by gas pressure changes due to neutral heating) will also induce ion drifts (or changes in ion drifts) parallel to the local magnetic field. Such 'parallel' ion drifts will also induce a field-aligned electron flow, to maintain quasi-charge neutrality. Thus the entire vertical plasma distribution will respond to wind changes, an effect which becomes increasingly important at greater altitudes. This change of ion density distribution will modify the consequent ion drag on the neutrals, and thus the wind acceleration terms, and finally the winds themselves. Since, in the vicinity of the auroral oval and polar cap, there are always various contra-flowing streams of field-aligned thermal and supra-thermal particles, it is difficult to identify the net FAC by direct observation. Yet it is the net flow which powers the magnetosphere - thermosphere forcing process, and variations of this field-aligned flow caused by feedback processes are important, but necessarily second order changes, and thus difficult to observe directly.

Large changes of upper thermospheric composition are induced by combined advection and convection, resulting from strong thermospheric heating in the polar regions. These composition changes (and specifically, enhanced concentrations of molecular nitrogen) are particularly pronounced at F-region heights during geomagnetic storm periods and within the summer geomagnetic polar region. Enhanced concentrations of molecular nitrogen cause significant depletions of F-region plasma densities by greatly increasing the effective recombination coefficient, while the ionisation rates, due to the combination of solar photoionisation and auroral precipitation, are only slightly changed.

At E-region altitudes, induced neutral wind changes, particularly those caused by ion drag, generally oppose the external magnetospheric EMF. This decreases ion drag acceleration (limiting the maximum induced winds), the local electrojet current (at all heights) and the Joule heating, since the electromotive force e.$(\underline{E} + \underline{V}_n \wedge \underline{B})$ is decreased. These processes are independent of any plasma density and conductivity modifications, however, the thermosphere will respond to the reduction in both electromotive force and electrojet current and consequent Joule heating.

Any modifications of the horizontal current (usually decreases) due to the induced winds, affects the capacity of the ionosphere to carry FAC connecting to the magnetosphere. Intuitively, the feedback effects of the induced winds on the total electromotive force, and the modified capacity of the auroral ionosphere to transmit or connect the FAC might be expected to cause some significant effects on the magnetosphere at times of large disturbances, when the E-region winds are known to reach 50 % of the $\underline{E} \times \underline{B}$ ion drift

velocity, driven by magnetospheric electric fields /20-22/.

Although some numerical experiments in these areas are in progress /23/, theoretical and experimental exploration of these problems is still at a very preliminary phase. We do not understand whether limits of the availability of charge carriers (ionospheric or magnetospheric) are important. Slight changes of the FAC may be matched by compensating changes in field-aligned potentials. Alternatively, slight shifts may occur in the patterns of overall magnetospheric convection, or its mapping to the ionosphere, to compensate for thermospheric or ionospheric feedback processes /24/. Feedback processes within the thermosphere and ionosphere probably affect a number of magnetospheric processes which have, historically, been thought of as purely magnetospheric / plasma physics phenomena.

The processes by which interactions between the polar ionosphere, thermosphere and magnetosphere occur, and the major routes by which feedback can occur, are easy to describe intuitively. However, the thermosphere and ionosphere are so closely interwoven by chemistry, dynamics and energetics, that quantitative modelling cannot rely on an intuitive approach.

The initial versions of the global 3-dimensional time-dependent numerical thermospheric models /7-11/ used theoretical models or the simple empirical Chiu /25/ global model of the ionosphere to calculate ion drag and Joule heating. However, the lack of any response at high latitudes to geomagnetic processes (precipitation, convection) within the Chiu ionospheric model caused a gross underestimate of the magnitude of ion drag and of Joule / frictional heating at E-region altitudes in the auroral oval. The F-region plasma densities of the Chiu model did not so seriously underestimate reality. When the Chiu model was used in the 3-D T-D (or GCM) models, it was possible to simulate F-region winds and temperatures within the upper thermosphere which were realistic for quiet conditions. However, under disturbed conditions, and in the E-region at all times, the Chiu /25/ ionospheric model simulated winds and currents which were unreasonably low /13-15/.

The underestimate of E-region momentum forcing and heating also caused very serious problems when simulating the geomagnetic response in neutral chemistry and in total density, as well as the E-region winds etc /14,15/. Thus a series of numerical experiments were started to find methods of coping more realistically with the underlying difficulty: how to simulate the mutual interactions of the polar ionosphere and thermosphere.

The first serious efforts to create an interactive model for the polar ionosphere and thermosphere were published by Quegan et al /26/ and Fuller-Rowell et al /27/. In this model, data sets from the UCL global thermosphere and the 'Sheffield' polar ionosphere model (UT-independent) were iteratively exchanged until stability was achieved. The effects of the model iterations showed that significant changes in plasma density were caused by the effects of induced winds. The auroral oval plasma densities were greatly enhanced compared with those of the Chiu model. As a result, induced thermospheric winds and heating were generally greatly increased compared with previous simulations using the global Chiu model.

This 'simple' coupled model could not, however, be universally applied to study UT variations, let alone the effects of variable solar, geomagnetic and seasonal conditions. The next stage, developing a fully interactive thermosphere and polar ionosphere model, was complex, and is described in some detail in a series of papers /28-31/. This fully coupled model, exchanges ionospheric and thermospheric parameters throughout the region poleward of $45^{o}$ geomagnetic latitude. At lower latitudes, the numerical model is presently still dependent on empirical ionospheric descriptions. When the physics of the major ionospheric-thermospheric interactions, are included within the coupled model, many of the additional 'geomagnetic' energy sources required previously to explain observations, are unnecessary /29/.

The approach taken in the development of the coupled model described above differs from that employed by the Utah State group /32/. In the latter modelling, the structure, dynamics and composition of the neutral atmosphere is assumed to be invariant to the response of the ionosphere to solar and geomagnetic forcing. In this work, we show that induced neutral atmosphere winds, temperature and composition changes do, in fact, cause major feedback changes of the ionosphere.

## 2. THE COUPLED GLOBAL THERMOSPHERE / POLAR IONOSPHERE MODEL.

The development of the UCL Three Dimensional Thermospheric Model is well documented in previous publications /7,8/, as is the Sheffield ionospheric model /26,33/. Progress with this model coupling has been reported in a number of papers during the past 3 - 4 years 26-31/.

The UCL Three Dimensional Thermospheric Model (or GCM) simulates the time-dependent structure of the vector wind, temperature, density and composition of the neutral atmosphere, by numerically solving the non-linear equations of momentum, energy and continuity /7/, and a time-dependent mean mass equation /8/. The global atmosphere is divided into a series of elements in geographic latitude, longitude and pressure. Each grid point rotates with the earth to define a non-inertial frame of reference in a spherical polar coordinate system. The latitude resolution is $2^{o}$, the longitude resolution is $18^{o}$, and each longitude slice sweeps through all local times, with a 1 min time step. In the vertical direction the atmosphere is divided into 15 levels in log (pressure), each layer is equivalent to one scale height thickness, from a lower boundary of 1 Pascal at 80km height.

The time-dependent variables of southward and eastward neutral wind, total energy density, and mean molecular mass are evaluated at each grid point by an explicit time stepping numerical technique. After each iteration the vertical wind is derived, together with temperature, heights of pressure surfaces, density, and atomic oxygen and molecular nitrogen concentrations. The data can be interpolated to fixed heights for comparison with experimental data, or with empirical models. The momentum equation is non-linear and the solutions fully describe the horizontal and vertical advection, i.e. the transport of momentum.

The neutral atmosphere numerical model uses an Eulerian approach. However, the ionospheric code /26-29/ has to be evaluated in a Lagrangian system. The complex convection patterns imposed by a magnetospheric electric field on plasma movements within the polar regions are referenced to a fixed Sun - Earth frame, assuming pure $\underline{E} \times \underline{B}$ drifts. The electric field is derived by merging a model of magnetospheric convection /4/ with the co-rotation potential (induced by the earth's rotation). Parcels of plasma are traced along their convections paths, which are often complex.

In the ionospheric code, atomic ($H^+$ and $O^+$) and molecular ion concentrations are evaluated over the height range from 100 to 1500 km, and used in the thermospheric code within $45^o$ magnetic latitude of the north and south magnetic poles. The use of the self-consistent ionosphere at high-latitudes and an empirical description at low- and mid-latitudes can result in a discontinuity in the vicinity of the plasmapause. The ionospheric code is presently being extended to include the self-consistent calculation at mid- and low-latitudes, including computation of the equatorial anomaly, but these new results will not be discussed here.

To avoid further repetition of the details of the numerical codes, the reader is referred to the series of papers previously cited for details of the UCL thermospheric model /7,8/, for information on the Sheffield ionospheric code /26,33-35/ and for information on the procedures used to couple the UCL thermospheric and Sheffield polar ionospheric models /28,29/.

## 3. COUPLED MODEL SIMULATIONS.

Two simulations of the coupled thermosphere and ionosphere for December solstice and two for the June solstice, using the UCL / Sheffield coupled model will be presented. These simulations have been generated for IMF BZ negative (southward), and for conditions when the IMF BY component was either strongly positive, or strongly negative, for a geomagnetic activity level corresponding to approximately Kp = 3 to 4, and for moderately high solar activity ($F_{10.7}$ cm = 185). The simulations are time-dependent, that is they are UT dependent, and the results are diurnally reproducible. However, the external solar and geomagnetic inputs are time-independent. These two simulations use an offset dipole representation of the geomagnetic field.

The characteristic UT variations of the summer and winter polar regions are dependent on the offset of the geomagnetic poles from the geographic poles. During the UT day, at all seasons, the geomagnetic polar caps are carried into and out of sunlight. There is, therefore, a large diurnal modulation of the solar photoionisation and UV / EUV heating of the geomagnetic polar regions which also causes large UT variations in plasma density, conductivity, ion drag and Joule and solar heating of the polar thermosphere. There are consequent large UT modulations of the thermospheric and ionospheric response. These characteristic UT variations of thermospheric and ionospheric structures, and the associated thermospheric-ionospheric interactions are discussed in /29/.

Figure 1 illustrates the global wind and temperature structure for the December solstice, at pressure level 12, close to 300 km altitude, at 18 Universal Time (UT) and moderate geomagnetic activity ($K_p$=3$^+$). A large temperature gradient has developed from the summer to winter hemisphere of over 500 K, due to the differential heating by the solar UV and EUV radiation. Superimposed on this latitude structure is a smaller amplitude diurnal variation of about 250 K range at equatorial latitudes. At high latitudes the magnetospheric momentum and energy sources resulting from the convection electric field, auroral particle precipitation and Joule heating have increased the neutral wind velocities, with the additional heat source coming primarily from Joule heating. In the winter hemisphere this additional heating from Joule and auroral particle sources reverses the global summer to winter temperature gradient, so that minimum temperatures occur at winter mid-latitudes.

The global circulation pattern responds to the large-scale temperature and pressure fields, creating a prevailing summer to winter flow at mid and high altitudes, which is particularly pronounced under quiet geomagnetic conditions. For average or high geomagnetic activity, the geomagnetic heating in the summer hemisphere enhances this seasonal flow, whereas in winter it opposes the solar forcing. A mean wind then flows from the winter pole to winter mid latitudes. The global pattern of mean molecular mass in the upper thermosphere is largely a result of these large-scale horizontal, and associated vertical wind fields. High values of mean molecular mass (above 20 at F-region altitudes) are found over the summer geomagnetic polar region, even under quiet conditions. As the geomagnetic activity rises, the summer polar values increase further, as high as 24-25, while values of above 20 can be found over the winter geomagnetic pole under active conditions. At high winter mid- and low-latitudes, the mean molecular mass is normally depressed, close to 16, nearly pure atomic oxygen.

Figure 2 illustrates the equivalent global structure of wind and temperature in the lower thermosphere, at pressure level 7 around 125 km, and again at 18 UT. At mid and low latitudes a large part of the response is due to the influence of semi-diurnal tidal modes propagating from the middle and lower atmosphere. At high latitudes, for this case of moderately disturbed geomagnetic activity, the wind pattern is controlled more by the electrodynamic forcing from the magnetosphere, than either in-situ solar forcing or propagating tides. At high latitudes the lower thermosphere wind magnitudes are about a factor two smaller than those shown in Figure 1 for the upper thermosphere. Note that the wind scale differs by a factor of two compared with the previous figure. However, it is interesting that the polar wind patterns are actually remarkably similar. The ion drag momentum forcing is different, not only due to reduced velocities, but also due to rotation of the ion drift velocity vector toward the electric field vector direction. However, compensation by the pressure gradient, Coriolis and advection terms (mainly) causes the steady-state wind pattern to follow that at higher altitudes, where ions follow the $\underline{E}\times\underline{B}$ direction. At mid and low latitudes the influence of the propagating tides has brought the magnitudes of the winds to levels similar to those seen in the upper thermosphere, of the order of 100 m/s. At the lower levels the winds are predominantly semi-diurnal, however, compared with a mainly diurnal variation in the upper levels, so that wind patterns of the upper and lower thermosphere are quite different.

## 4. POLAR F-REGION WINDS, COMPOSITION AND ION DENSITY.

Figure 3 illustrates the response of the upper thermosphere wind and composition, the mean molecular mass, in the northern polar regions (50-90$^{o}$) to the orientation of the IMF-BY in summer and winter. Figure 3a and b are for winter, IMF-BY negative and positive respectively, 3c and d are the equivalent for the summer.

At F-region altitudes, only relatively small differences in the wind patterns can be ascribed to seasonal effects. In the winter, the winds in the dawn auroral oval show only a slight tendency to follow the relatively weak sunward ion convection (BY positive). In the summer, there is no significant indication of sunward wind acceleration at all in the dawn cell for BY positive. Within the strong clockwise circulation wind cell which follows the strong sunward ion convection in the dusk auroral oval and antisunward ion flow over the dawn side of the polar cap, only minor seasonal differences occur. Equatorward of the nightside auroral oval, the wind flow is distinctly equatorward in the June simulation (200 m/sec), and rather weaker in the December simulation (100 m/sec).

There are two major responses of high-latitude F-region winds to changes of IMF BY sense. There is shift of a region of high-velocity anti-sunward flow from the dusk (BY negative) to dawn (BY positive) side of the polar cap. This has been previously reported /15,36/ in observations from the DE-2 spacecraft and from the ground. The winds of the dawn auroral oval also show a characteristic change. For IMF BY positive, there is very little indication of sunward wind flow in response to the weak dawn auroral oval / polar cap convection cell. For IMF BY negative there are significant sunward winds in the dawn auroral oval. However, these are always weaker than the strong sunward winds of the dusk auroral oval. The primary reasons for this dusk / dawn asymmetry have been discussed previously /27,37/. The asymmetric response is caused by a natural atmospheric 'resonance' responding to the clockwise wind vortex forced by sunward ion convection in the dusk auroral oval and antisunward ion convection over the polar cap, given the sense and rate of rotation of the Earth.

Coriolis and curvature accelerations balance within the clockwise vortex, of which the dusk auroral oval is part. In the dawn auroral oval, the conditions for this resonance do not exist. As a result, for the same plasma densities and convection velocities, there is a much smaller effective sunward wind acceleration. This effect was noted in earlier rocket wind measurements /21/, and has been well observed by DE-2 /12,13/. The effect was predicted by much simpler simulations than those we are describing here /7/. However, it is perhaps reassuring to see that the asymmetric dusk / dawn wind response is still observed in the present rather more sophisticated simulations, where the possibility of a complex feedback process between the thermosphere and ionosphere, which might change the nature or magnitude of the wind acceleration and response can now be discounted, since the coupled model accounts for all the major feedback mechanisms with the possible exception of feedback processes involving magnetosphere / ionosphere coupling.

The seasonal variation of neutral thermospheric composition at fixed height level is well established by experimental observations /38,39/. At mid latitudes, enhanced concentrations of $[N_2]$ relative to [O] in the summer hemisphere between 250 and 350 km feeds back into the ionospheric chemistry by increasing the recombination rate of the dominant ion $O^+$, decreasing the $N_mF_2$ (peak electron density at the peak of the $F_2$ layer). Conversely, in the winter hemisphere, the enhanced $[O/N_2]$ ratio decreases the recombination rate of the F region $O^+$ ions, and causes a general increase in $N_mF_2$, despite the overall decrease of solar insolation.

These general features are well illustrated by comparisons between Figures 3a and 3c, for the F-region (IMF BY negative) or Figures 3b and 3d (IMF BY positive). The highest values of mean molecular mass which occur poleward of 50$^{o}$ latitude in the winter polar F-region equal the lowest value in the equivalent summer polar region. At high winter mid-latitudes, the mean molecular mass is close to 16, indicating a composition which is nearly pure atomic oxygen. There is then a winter-time plateau around the geomagnetic pole, due to heating, upwelling and outflow, where the mean molecular mass reaches 20.

This value of mean molecular mass within the winter plateau is very dependent on geomagnetic activity. At very low activity levels the mean molecular mass may be as low as 16 - 17. During extended geomagnetic storm conditions, the mean molecular mass may reach values as high as 22.

In the summer polar cap, the lowest values of mean molecular mass above $50^{o}$ are 20, and the highest values within the geomagnetic polar cap reach 24 to 25. At this constant pressure level (12), this implies a 4 fold reduction in the density of atomic oxygen between high summer mid-latitudes and the pole. Atomic oxygen concentrations at the summer pole are a factor of 10 lower than those found at high winter mid-latitudes. The variation of molecular nitrogen density is in direct anti-phase, and compensates for the atomic oxygen changes.

There are significant changes in the compositional variations caused by changes in the sense of IMF BY. These detailed changes in mean molecular mass are modest (1 - 2 units) compared with the larger changes (3 - 4 units) induced by the combination of seasonal solar insolation changes and the heating and consequent overturning of the thermosphere resulting from geomagnetic forcing. However, the F-region plasma density is highly responsive to the combination of BY-induced modulation of neutral composition and ion transport.

Figure 4 illustrates the response of the F-region plasma density for the same conditions as described in Figure 3. Comparing Figures 4a and 4c, (IMF BY -ve) or Figure 3b and 3d (IMF BY +ve), the high-latitude seasonal F-region ionospheric anomaly is apparent: in some areas the winter high latitude plasma densities exceed those of the equivalent summer, sunlit, region. Peak densities in the winter sunlit cusp region (which is just in sunlight at 18 UT) are about a factor of 2 - 3 higher than the equivalent summer region. The mean levels over the high latitude region are similar in summer and winter at this UT.

The winter polar cap, at 18 UT, is filled with high density plasma which is rapidly convected from the dayside through the cusp region. There is a well-developed sub-auroral trough, from 16 LST to 08 LST at this UT (but which encircles the entire auroral oval at the winter solstice for much of the rest of the UT day). At UT times when the cusp is not in sunlight, the plasma density within the polar cap tends to fall to significantly lower values /29/.

In contrast, the summer sunlit polar cap is a region where the electron recombination rate is high, so that plasma densities are low, despite the combination of solar photoionisation and auroral electron precipitation. Regions of lowest plasma density have a one-to-one relationship with regions of the highest mean molecular mass. The regions of lowest plasma density within the summer polar cap are responsive to thermospheric compositional changes, induced by convection and Joule heating. The details of these responses are highly sensitive to the convection response to IMF BY changes, and these will be discussed later.

In the summer geomagnetic polar region there is no feature which corresponds to the winter sub-auroral trough There is sufficient solar photoionisation in the sub-auroral regions, which are regions of plasma stagnation, at least at 18 UT, that a sub-auroral trough does not develop. The equivalent winter sub-auroral regions are where the stagnation troughs do develop. There is enhanced plasma destruction compared with the equivalent winter sub-auroral region, but the values of mean molecular mass (20 - 21) are not high enough to cause the extreme plasma destruction rates which occur within the summer geomagnetic polar cap.

There is one caveat concerning this overall conclusion related to the southern summer polar region. At Universal Times when the geomagnetic polar cap is rotated furthest from the dayside (03.6 UT), and due to the greater offset of the southern geomagnetic pole from the geographic pole, there is no solar photoionisation on the nightside in the sub-auroral oval. A partial summer sub-auroral trough may occur under such conditions. This feature will, however, never develop to the extent seen in the winter hemisphere. For most of the UT day, in winter, the sub-auroral trough completely or nearly completely encircles the northern auroral oval. The situation at 18 UT, depicted for all these simulations, is when the winter northern cusp is sunlit, however, this only occurs for about 6 hours around 18 UT.

Sub-auroral troughs which develop in the summer hemisphere are always likely to be the fossils of an earlier, expanded, auroral oval. Intense heating, forcing enrichment of molecular species by convective overturning and advection will lead to the rapid destruction of plasma once the auroral source diminishes at the end of the disturbance. This is likely to be the mechanism creating such troughs, which is therefore quite distinct from the plasma stagnation mechanism which is mainly responsible for the winter sub-auroral trough.

In a comparison between Figures 4a (BY negative) and 4b (BY positive) the major response of the winter F-region plasma density in response to IMF BY changes can be seen. For BY negative (4a), a long extended tongue or plume of plasma is rapidly transported antisunward from the dayside and cusp. This rapid transport is produced by the situation that the strong antisunward plasma convection on the dusk side of the polar cap and the effect of co-rotation in that location are in the same direction. On the dawn side of the polar cap, the weak antisunward plasma convection is opposed by co-rotation so that a deep plasma hole develops (no sunlight, little precipitation).

When IMF BY is positive, the rapid antisunward plasma transport on the dawn side of the polar cap is opposed by co-rotation. The corresponding tongue or plume of anti-sunward moving plasma is thus weaker and shorter, and tends to lie in the central polar cap, rather than the dawn side. The co-rotation velocity is not opposed to convection in the central polar cap region. Co-rotation also aids the otherwise weak antisunward transport on the dusk side of the polar cap, another effect which broadens the width of the antisunward plasma tongue or plume.

The effect of combined convection and co-rotation for IMF BY negative is sufficiently strong that for the same precipitation, the dusk auroral oval plasma densities are roughly double those for IMF BY positive. This is entirely due to the shorter time required to transport plasma from the dayside cusp through the polar cap and into the dusk auroral oval for IMF BY negative.

There are relatively few changes of the winter sub-auroral trough which are produced by the response to IMF BY. For IMF BY positive, the dusk trough is somewhat broader in latitudinal extent than for BY negative, due to the stronger sunward convection in the dusk oval, but the effect is relatively minor, and must be dependent on the coincidence (or non-coincidence) of convection and precipitation boundaries.

The plasma density structures and values within the winter auroral oval do not appear to be strongly dependent on IMF BY. In the general sense, there is rather little enhancement of F-region plasma density associated with the auroral oval, although a range of characteristic signatures in ion and electron temperature, as well as changes of ion, and neutral composition are produced.

In the summer polar cap F-region, the differences are quite dramatic. Deep troughs form in regions where there is strong photoionisation and some particle precipitation, caused by the neutral compositional response to strong heating: the stronger the ion convection, the stronger the ion heating and, generally, the stronger the heating of the neutral atmosphere. The rapid convective upwelling, combined with horizontal advection causes the very large increase in F-region mean molecular mass, which then causes the plasma troughs.

The most important conclusion which comes from this study, related to F-region behaviour and ionosphere-thermosphere coupling, is the distinctive seasonal behaviour. In the winter polar region, it is primarily plasma convection which determines the distribution of plasma density and so modulates the local frictional heating and the acceleration of winds by ion drag. The transport of F-region plasma by convection and the combination of convection and co-rotation control the existence and location of tongues of high plasma density within the polar cap and troughs of low plasma density within the polar cap and in the sub-auroral trough. Neutral wind effects are significant through the vertical transport of plasma. The convective overturning of the neutral atmosphere and associated advection causes significant compositional changes which cause noticeable, but not dominant effects on the winter-time F-region plasma distribution.

In the summer hemisphere, the effects of combined solar insolation and geomagnetic heating force a much greater convective / advective overturning of the thermosphere. In the summer polar cap, at levels of moderate geomagnetic activity, the greatly increased proportion of molecular species causes a fundamental change in the F-region ionospheric chemistry. As a result, compositional changes, driven by the detailed response of the thermosphere to geomagnetic forcing, dominate the detailed ionospheric response. Regions where, in the winter pole, the plasma density is enhanced by rapid convection from the dayside, tend to become regions of low plasma density in summer. The combination of plasma convection and co-rotation are still important. However, in the summer polar cap, the plasma density distribution is controlled by the neutral composition distribution. The composition is modified by two processes: changes in the location of Joule heating and consequent upwelling, and horizontal transport by the neutral wind. Plasma depletions are signatures of an increase in mean molecular mass.

While in the course of this paper, we have only considered two, asymmetric, convection patterns, the general conclusions described above are true for other representations of the convection field /5/. Detailed inter-comparisons with those described in the present study confirm the conclusions of the previous two paragraphs /29/.

## 5. ION TEMPERATURE AND ION COMPOSITION.

Figures 5 and 6 illustrate the ion temperature, $O^+$, molecular ion, and $H^+$ number density distributions for the winter and summer northern polar regions, respectively. Both figures are from the simulations with IMF-BY negative, at 18 UT, and the altitude is 301 km. For the solar activity depicted here ($F_{10.7}$ = 185), this altitude is close to the F2 peak in winter but significantly below it for the summer case.

The patterns of ion temperature are virtually identical in summer and winter, the structure being highly dependent on the magnitude and direction of the ion drift vectors, which in turn are dependent on the orientation of the IMF-BY which, of course, is the same in these two figures. The main difference is in the magnitude, the summer value exceeding the winter by about 500 to 600 K, approximately equal to the difference in the neutral temperature. The influence of the neutral wind will not be addressed specifically in this paper. Obviously, the patterns of ion temperature will be convection-pattern dependent.

A further consequence of the global seasonal composition difference is the effect on the relative proportion of atomic and molecular ions in the polar regions. The patterns of ion composition are substantially different in the two seasons. This contrast can be seen in Figures 5 and 6. The general magnitudes of the $O^+$ number density in the polar regions are very similar in summer and winter at this altitude, although the reasons for their individual structure are very different. The winter polar F-region, although in darkness, is maintained by the transport of $O^+$ from the dayside sunlit sub-auroral ionosphere where the winter plasma densities are consistently larger than corresponding summertime values. The very slow recombination rates in the neutral oxygen-rich atmosphere is sufficient to keep the winter polar $O^+$ density above the summer values. In fact the dominant ion is $O^+$ in the winter at these altitudes, the mechanisms creating the pattern are therefore identical to the previous discussion for the total ion density. The molecular ions in winter at this altitude contribute generally less than 5% to the total ion density. The pattern of molecular ions, predominantly $NO^+$, has a similar trend to that of the pattern of ion temperature (and Joule heating). The regions of upwelling are related to regions of enhanced molecular ions.

In summer the ratio of atomic to molecular ions is very much reduced. In fact, even at this altitude of 301 km, some regions of the polar thermosphere in summer have more molecular than atomic ions. Generally the levels are very similar over much of the summer polar regions. Outside the polar plateau of enhanced molecular ions, associated with the similar plateau in the neutral thermosphere, the proportion of molecular to atomic ions decreases rapidly toward mid-latitudes.

The $H^+$ number density is typically a few orders of magnitude less in concentration than the total ion density, although it does become significant at much higher altitudes, in excess of 500 km.

## 6. LOWER THERMOSPHERE WINDS AND ION DENSITY.

Figure 7 illustrates the neutral wind vector and plasma density in the lower thermosphere, at pressure level 7, about 225 km altitude, close to the peak in the Pedersen conductivity. The figure shows the same four conditions described in Figures 3 and 4. The major changes of winds of the E-region result from the distinctly different pole to pole thermal structure at the June and December solstices. There is a pronounced cyclonic vortex around the winter polar region, which is observable at high mid-latitudes, superimposed on a generally weak flow away from the sub-solar point. This circulation is replaced by a pronounced anti-cyclonic vortex around the summer polar region (again best seen at high mid-latitudes).

The wind patterns within the auroral oval and polar cap show little change with season for the same convection and precipitation patterns. The major wind features result from ion drag, and are thus responsive to convection changes and plasma density enhancements (resulting from combined particle precipitation and solar photoionisation). The basic auroral oval wind pattern follows that of ion convection, as in the F-region. The E-region wind velocities are typically about a factor of 2 - 3 lower than at F-region altitudes.

A highly satistactory feature of the present self-consistent simulations is that they show E-region winds, at times of moderate geomagnetic disturbances, which are of similar magnitudes and structures to those reported in experimental studies of E-region winds /20-22/. Until the present self-consistent thermosphere - ionosphere code was developed, such correspondence of simulated E-region winds with the real world could only be obtained by extensive and artificial tuning of E-region plasma density distributions (even in the earlier code /26,27/).

Seasonal changes of E-region winds are relatively small, however, the response to IMF / plasma convection changes is very marked. Obviously, this response is also modulated by the precipitation, which causes the large increases of ion drag coupling via enhanced E-region plasma density enhancements. In general, therefore, polar E-region winds will be highly dependent on actual convection patterns.

For IMF BY positive, there is a large clockwise swirl, sunward around the dusk auroral oval and antisunward over the dawn side of the polar cap. The peak E-region wind speeds in the dusk auroral oval reach 200 - 250 m/sec, which is quite consistent with observations /40/. There is a relatively weak sunward wind flow in the dawn auroral oval (less than 50 m/sec), combined with a significant equatorward flow component (50 - 100 m/sec).

For IMF BY negative, the sunward wind flow in the dusk auroral oval is still strong, except that it only extends to about 14 LST. The antisunward flow over the polar cap is more generally spread over the entire width of the polar cap, rather than being confined to the dusk side, as tends to happen in the F-region (see earlier discussion). The sunward winds of the dawn auroral oval, for IMF BY negative are significantly increased, to nearly 100 m/sec, and there is a generally more coherent flow, following the ion convection of the dawn auroral oval.

The asymmetry of the wind response in the dusk and dawn parts of the auroral oval is again related to the conditions required for resonance /37/. As noted previously, the winds follow the $\underline{E} \wedge \underline{B}$ direction more than would be expected, considering the rotation of the ion drift vector toward the $\underline{E}$ (electric field) direction at these altitudes. This is the result of the stable, steady-state, response of the atmosphere. A short-period impulse would have rather different consequences, with the initial winds following more closely the mean ion vectors, and later responding to changes in pressure gradients, Coriolis, advection and

viscosity terms. Clearly, again, the detailed response will be convection-field dependent.

In the F-region, peak values of mean molecular mass normally are associated with the highest neutral temperatures. In the E-region, although there is a general polar plateau of high mean molecular mass, the fine details of the thermal and compositional distributions are not so closely associated. For example, there are consistent peaks of mean molecular mass on the dawn side of the polar cap, in a region where the E-region neutral temperature tends to be rather low. This feature, both in temperature and in composition, is due to the anti-clockwise (cyclonic) ion flow vortex. The curvature and Coriolis acceleration terms /37/ have the same direction in this region and the neutral gas is spun outward, in addition to the tendency for induction of anticlockwise wind flow by ion drag. The induced outflow causes enhanced upwelling, yet another example of the complexity of the variety of interactions between the polar ionosphere and thermosphere, and the respective roles of plasma convection, precipitation, and the neutral gas response.

Under all the conditions of these four simulations, the E-region ionospheric chemistry is not fundamentally changed, despite the seasonal modulation of mean molecular mass described earlier. The plasma density distribution in regions outside the auroral oval shows an increase (as expected) in the summer hemisphere, as the result of increased photoionisation. E-region plasma density values on the nightside in the winter hemisphere are predicted to be low (2 - 5 $x10^9$) in places. Within the geomagnetic polar cap, plasma densities within the winter hemisphere are lower than in the summer hemisphere. This is the direct result of additional solar photoionisation within the summer polar cap. On average, the E-region values within the winter polar cap are about a factor of 4 lower than in the summer polar cap. The additional summertime ionisation increases Joule heating and ion drag forcing of the neutral atmosphere.

What we see from this study is that E-region plasma densities around the auroral oval are primarily enhanced as the result of precipitation. The seasonal modulation and the effects of plasma wind transport are generally of little importance. Vertical plasma transport due to the horizontal winds is likely to be the most important single effect, and this may be particularly important in connection with the formation and subsequent behaviour of sporadic E layers, resulting from metallic ions. Such long-lived ions are not subject to the rapid destruction of the E-region molecular ions which alone are considered in this study. For metallic E-region ions, horizontal as well as vertical transport effects will be much more significant /41/.

Related to the rapid recombination rates typical of molecular ions at E-region altitudes, there are few significant features of the E-region plasma distribution which can be attributed to convection, and thus to IMF BY changes. This would not be the situation for longer-lived ions, such as the metallic species found in sporadic E layers. However, with the exception of such localised and usually ephemeral species and features, the present simulations indicate that primarily, the E-region plasma distribution is controlled by the combination of solar photoionisation and energetic electron precipitation. The plasma density enhancements are a critical factor in producing observed winds, and conductivity enhancements.

The actual values of ion drag wind acceleration terms, ionospheric conductivities, electric currents, and the implicit distribution of magnetospheric FAC, which power the entire high-latitude system, are therefore determined primarily by precipitation: Typical peak E-region ion density values within the moderately disturbed auroral oval are a factor of 3 greater than those of the surrounding sunlit, but non-auroral, ionosphere in summer time. In winter time, the corresponding enhancement factor due to electron precipitation is between a factor of 10 and 100 times above the non-sunlit E-region ionosphere.

## 7. ELECTRODYNAMICS.

In the following section, two of the electrodynamic parameters will be discussed and compared from the four numerical simulations. These parameters are the height-integrated Joule heating rate and the FAC system. Unlike the previous figures which have been referenced to a fixed pressure or height level, these new parameters are a consequence of the integrated response of the thermosphere and ionosphere at all the levels. Generally both these quantities have their maximum contribution from the lower thermosphere, but there are times and locations when this is not the case, and a significant contribution can come from the upper levels.

The height integrated Joule heating rate is a sensitive parameter. This is due to its dependence on the second power of the electric field, and also due to its critical dependence on the colocation of boundaries of conductivity enhancements through electron precipitation and maximum ion drifts. Obviously, one must treat with caution the idea of using the product of two independently produced models, particularly when one expects a strong causal relationship between the two sources via their magnetospheric origin. Ideally, the models of precipitation and convection should be assembled 'self-consistently'. Any correlation or anti-correlations (as have been reported) between the two data sets can then be maintained. However, no convenient empirical of statistical model of the combined magnetospheric sources is available, so we must rely on using, with some caution, those currently at our disposal.

Figure 8 illustrates the height-integrated Joule heating rate for the same four cases as presented in Figures 3, 4 and 7. The shape and magnitude of the patterns is very strongly dependent both on season and the orientation of the IMF-BY (and thus on any specific

convection field). Due to the second power dependence on the neutral wind - ion velocity difference, the peak Joule heating rates tend to be associated with the peaks in the ion drift vectors. A particular characteristic of using electric field models with two extreme orientations of the IMF-BY, is that either the dusk or dawn boundaries between polar cap and auroral oval have a double-maximum feature in Joule heating. These double features are created by the combination of the expected peak that occurs in the dawn and dusk sector auroral oval (maximum sunward ion drift), and a second feature associated with the enhanced ion flow over either the dusk or dawn side of the polar cap. For IMF-BY negative the twin maximum feature is on the dusk side, for IMF-BY positive it is on the dawn side.

The F-region ion densities in the polar regions were similar in magnitude in summer and winter. In the lower thermosphere the situation is more diverse. Within these numerical simulations, neither the precipitating electron flux nor the electric field has a seasonal dependence. Within the auroral oval therefore, where the main contribution to the Pedersen conductivity arises from the auroral precipitation, the peak Joule heating rates do not have a strong seasonal dependence.

Outside the auroral oval the situation is different for two reasons. Firstly, there is a large difference in the lower thermosphere conductivity in summer and winter. Secondly, in winter the lower thermosphere conductivity is very small, so that a significant contribution to the height-integrated conductivity, and consequently also the Joule heating rate, comes from the upper thermosphere. As was shown before, the winter F-region maintains as much if not more plasma concentration, compared with summer values. The peak values outside the auroral oval reflect these seasonal differences in the conductivity of the thermosphere. As a result, the hemispheric integrated Joule heating rates maximise in the summer hemisphere.

The seasonal difference is most clearly seen by comparing Figure 8a and c, for winter and summer conditions respectively. In the auroral oval the dusk or dawn auroral oval values have changed by about 50% due to the small relative change in ionization from solar radiation. The polar cap feature however, in this case on the dusk side, is considerably increased, in some areas increasing in magnitude by more than a factor of three. The extreme difference in magnitude of this polar cap feature is due to the fact that in winter, at this time, the location of the peak associated with the fast ion velocities near 0200 LT is just in darkness, and experiences modest fluxes of auroral precipitation. In summer this same location now has a zenith angle close to $50^{o}$, at this time, and the change in solar ionization dominates the response.

Peak Joule heating values for both simulations with IMF-BY negative occur in the early dusk magnetic sector. In winter and summer, dusk auroral oval values exceed the dawn values by a factor of two. In summer the extreme polar cap maximum exceeds the dawn sector values by a factor of six.

For IMF-BY positive (Figure 8b and d) the seasonal differences are less dramatic, and both patterns are very different from Figures 8a and c. The double feature is now in the dusk sector, and the peak values are near midnight, within the auroral oval, and towards the dayside for the polar cap feature. The polar cap feature again relies on solar produced plasma, and again increases in magnitude by a factor of three. The midnight sector peak increases by about 50% from winter to summer.

The effect of ignoring the influence of the neutral wind on the height-integrated Joule heating rate can be seen in Figure 9, which shows data from the same four simulations. The dawn and midnight sector auroral oval values have changed little, but Joule heating in the dusk auroral oval has increased by 30 to 40%. This result is a direct reflection on the ability of the neutral wind to follow the ion motion. It has been discussed extensively both from observation /12/ and from simulations /27,42/ that the dusk sector neutral atmosphere is much more able to follow the ion drift vector. This effect has been explained as an inertial resonance effect, where the natural motion of the atmosphere at mid and high latitudes is to follow a clockwise vortex. The dusk sector ion motion naturally excites this trend driving the larger winds in this sector. Therefore when the $\underline{V} \times \underline{B}$ component is removed from the full expression for the Joule heating rate, the main effect is an increase in the dusk sector. In the polar cap regions of high Joule heating, on the dusk (BY negative) and dawn (BY positive) sides the peak values are reduced when neutral winds are excluded.

Deriving FAC from the statistical models of convection and precipitation, even using the fully-coupled thermosphere-ionosphere model, is stretching their relevance to the limits. The height-integrated horizontal ionospheric current system is obtained from a product of the conductivity and total electric field, and the FAC system is calculated from the divergence of the horizontal current. This condition then satisfies current continuity. Combining this calculation with the previously derived quantities, we then have an indication of the total demand made by the ionosphere on the magnetospheric sources of plasma, charge, momentum and energy. For the horizontal current to flow in the ionosphere, current continuity demands that either the magnetosphere is able to supply the necessary FAC, or polarization charges build up within the ionosphere, (or via instabilities triggered by large FAC or 'double-layers' etc), which tend to reduce the magnitude of the electric field imposed upon the upper atmosphere. The sensitivity of the magnetosphere to the feedback from the ionospheric current system is as yet difficult to quantify experimentally or theoretically.

Figure 10 show the FAC, for the northern polar regions from the same four simulations. Region 1 and 2 of the Birkeland (FAC) current system are clearly defined.

In winter the peak upward FAC are larger in magnitude than the downward FAC, with peak values a factor 1.5 to 2 larger. The regions of convergence of the height-integrated horizontal current which drive the upward FAC, are within the auroral oval. The regions of divergent horizontal current tend to lie outside those of auroral precipitation and therefore rely on solar-produced conductivities. The upward FAC is confined to a smaller area. Globally-integrated, the differences between total upward and downward FAC appear to be small, within physical and numerical uncertainties, as might be expected.

In summer the upward current increases only by about 50%. In contrast the downward current increases by a factor of 2-3, reflecting the importance of the increase in solar produced conductivities. The result is that in summer the peak in the upward and downward FAC roughly balance.

A number of differences occur in the FAC due to the orientation of the IMF-BY. The magnitudes for IMF-BY positive exceed those for BY negative by 50%. For BY negative the upward FAC closes on the dayside, to just form a continuous circle, surrounding the region of downward FAC. For IMF-BY positive the continuity of the upward current on the dayside is split by a region of downward current. The region of upward FAC is now diverted over the polar cap, but in the midnight sector the continuity of the pattern is much smoother

8. CONCLUSIONS.

Data are presented from a series of four numerical simulations of the coupled thermosphere and ionosphere. The data from high latitudes are shown at high (300km) and low (125km) thermospheric altitudes, to illustrate the response of the system to changes in the orientation of the IMF BY component at the summer and winter solstices. Each simulation is at high solar activity and moderate geomagnetic activity, and data are presented for 18 UT.

In the winter F-region, the plasma distribution is strongly controlled by transport. As the convection electric field moves into sunlight, solar produced ionization is 'picked-up' and carried over the polar cap. For By negative, a tongue of plasma extends over the dusk side of the polar cap, the convection electric field being assisted by the natural tendency of co-rotation. A polar ionisation hole is created by the lack of transport from regions of plasma generation, no photoionisation and negligible particle ionization. Plasma carried anti-sunward over the dusk side of the polar cap continues into the dusk auroral oval, where number densities exceed those in the dawn sector auroral oval by at least a factor of two For By positive the tendency of the convection electric field is to transport the plasma to the dawn side of the polar cap. The opposing influence of co-rotation in this case restricts the flow, resulting in a much smaller tongue over the central polar cap. Auroral oval ion densities in this case remain symmetric. In both cases a sub-auroral trough encircles the nightside.

There is also a characteristic neutral wind response in the upper thermosphere to the changing sense of IMF By. For By negative, enhanced anti-sunward winds occur on the dusk side of the polar cap, and modest sunward winds are generated in the dawn auroral oval. For By positive the peak anti-sunward winds move to the dawn side of the polar cap, and the sunward winds in the dawn auroral oval are virtually eliminated. In both cases substantial sunward winds occur in the dusk auroral oval, but are stronger for By negative due to the higher ion densities, resulting from fast ion transport from the dayside, across the dusk side of the polar cap and into the dusk auroral oval.

In summer the plasma distributions are flatter, are generally lower in concentration, and lack the sub-auroral trough, at least at 18 UT, as presented here. Transport is much less important in defining the plasma distribution which, in summer, at this level of geomagnetic activity, is controlled by the peaks and troughs in neutral composition. Plasma density minima correlate with peaks in the mean molecular mass, and contrast with the peaks in ion density that occurred at similar locations in winter. The generally higher mean molecular mass (i.e. enhanced molecular species) in summer is created by the global seasonal neutral circulation pattern. At high latitudes the pattern is modulated by regions of enhanced upwelling due to Joule heating, and the characteristic By-dependent composition distribution is finally imposed by the transport effect of the horizontal neutral wind. The neutral wind maintains its By dependence in summer, and changes very little with season.

The ion temperature in the F-region is higher in summer due to the seasonal difference in the temperature of the neutral atmosphere. The overall pattern of ion temperature is primarily dependent on the ion-neutral velocity difference, which will be a function of the IMF-BY direction, at least, and will depend little on season. Ion temperature distribution, as with many other features, will be highly dependent on the actual ion convection pattern.

The ratio of atomic to molecular ions at 300 km altitude has a strong seasonal dependence, virtually reflecting the strong seasonal difference in the neutral composition. In winter $O^+$ dominates, in summer the concentration of molecular ion increase dramatically, such that in some areas of the summer polar region they exceed the atomic ion concentration.

In the E-region, at this level of geomagnetic activity, the ion density in the auroral oval is dominated by auroral precipitation, which for these simulations has no By or seasonal dependence. Small changes do occur due to the seasonal changes in solar ionization. Outside the auroral oval, increased solar illumination in summer generates ion densities exceeding those in winter by about a factor of three.

The neutral wind in the lower thermosphere maintains the signature of the IMF By that was the hall-mark of the F-region circulation. It has a noticeable but not dominant seasonal variation due to the increase of ion density by the background solar illumination.

The height-integrated Joule heating rate shows a strong seasonal and BY dependence. A double-peak feature develops in the dusk sector for BY negative, and in the dawn sector for BY positive. Within the auroral oval, the seasonal increase is small, 50%, due to primary dependence of conductivity on auroral precipitation. In the polar cap the seasonal response is large, due to dependence on solar produced conductivity.

In the dusk auroral oval, Joule heating exceeds that in the dawn sector by a factor of two. Ignoring neutral winds enhances Joule heating, particularly in the dusk sector, thus increasing the ratio of Joule heating in the dusk/dawn sectors. The conductivity changes are also reflected in the patterns of FAC.

ACKNOWLEDGMENTS

We are indebted to Dr. Fred Rich for provision of the Heppner and Maynard polar electric fields in the form of harmonic coefficients. We would also like to express our particular thanks to John Harmer and Hilary Hughes for their assistance in preparing, running and processing the computer simulations using the UCL / Sheffield coupled ionospheric / thermospheric model. Computer time was made available by the University of London Computer Center (CRAY 1-S), and the CRAY-XMP-48 at the Rutherford Appleton Laboratory. The research was supported by grants from the UK Science and Engineering Research Council, and from the European Office of Aerospace Research and Development (AFOSR-86-341).

REFERENCES:

1. FULLER-ROWELL and D.S. EVANS, (1987), Height-Integrated Pedersen and Hall Conductivity Patterns Inferred from the NOAA/TIROS Satellite Data. J. Geophys. Res. 92, 7606-7618.
2. HARDY D., M.S. GUSSENHOVEN, E. HOLEMAN, (1985), A statistical model of auroral electron precipitation. J Geophys. Res., 90, 4229-4248.
3. HEPPNER J.P., (1977), Empirical Models of High Latitude Electric Field, J. Geophys. Res. 82, 1115-1125.
4. HEPPNER J.P. and N.C. MAYNARD, (1987), Empirical High-Latitude Electric Field Models, J. Geophys. Res. 92, 4467-4490.
5. FOSTER J.C., J.M. HOLT, R.G. MULGROVE and D.S. EVANS, (1986), Ionospheric Convection associated with Discrete Levels of Particle Precipitation, Geophys. Res. Lett., 13, 656-659.
6. TINSLEY B.A., Y. SAHAI, M.A. BIONDI and J.W. MERIWETHER, (1988), Equatorial particle precipitation during geomagnetic storms and relationship to equatorial thermospheric heating, J. Geophys. Res., 93, 270-276.
7. FULLER-ROWELL T.J. and D. REES, (1980), A three-dimensional, time-dependent, global model of the thermosphere, J. Atmos. Sci. 37 2545-2567.
8. FULLER-ROWELL T.J. and D. REES, (1983), Derivation of a conservative equation for mean molecular weight for a two constituent gas within a three-dimensional, time-dependent model of the thermosphere, Planet. Space Sci. 31, 1209-1222.
9. DICKINSON R.E., E.C. RIDLEY and R.G. ROBLE, (1981), A three-dimensional general circulation model of the thermosphere, J. Geophys. Res. 86, 1499-1512.
10. DICKINSON R.E., E.C. RIDLEY and R.G. ROBLE, (1984), Thermospheric general circulation with coupled dynamics and composition, J. Atmos. Terr. Phys. 41, 205-219.
11. ROBLE, R.G., R.E. DICKINSON and E.C. RIDLEY, (1982), The global circulation and temperature structure of the thermosphere with high-latitude plasma convection, J. Geophys. Res. 87, 1599-1614.
12. HAYS P.B., T.L KILLEEN. N.W. SPENCER, L.E. WHARTON, R.G. ROBLE, B.A. EMERY, T.J. FULLER-ROWELL, D. REES, L.A. FRANK, and J.D. CRAVEN, (1984), Observations of the dynamics of the polar thermosphere, J. Geophys. Res. 89, 5547-5612.
13. REES D., T.J. FULLER-ROWELL, R. GORDON, T.L. KILLEEN, P.B. HAYS, L.E. WHARTON and N.W. SPENCER, (1983), A comparison of the wind observations from the Dynamics Explorer satellite with the predictions of a global time-dependent model, Planet. Space Sci. 31, 1299-1314.
14. REES D., R. GORDON, T.J. FULLER-ROWELL, M.F. SMITH, G.R. CARIGNAN, T.L. KILLEEN, P.B. HAYS, N.W. SPENCER, (1985), The composition, structure, temperature and dynamics of the upper thermosphere in the polar regions during October to December 1981, Planet. Space. Sci. 33, 617-666.
15. REES D., T.J. FULLER-ROWELL, R. GORDON, M.F. SMITH, J.P. HEPPNER, N.C. MAYNARD, N.W. SPENCER, L.E. WHARTON , P.B. HAYS, and T.L. KILLEEN, (1986), A theoretical and empirical study of the response of the high-latitude thermosphere to the sense of the "Y" component of the interplanetary magnetic field, Planet. Space Sci. 34, 1-40.
16. FESEN C., R.G. ROBLE, E.C. RIDLEY, (1986), Simulations of thermospheric tides at equinox with the NCAR Thermospheric General Circulation Model, J. Geophys. Res., 91, 4471-4489.
17. RISHBETH H. and O.K. GARRIOT, (1969), Introduction to Ionospheric Physics, Academic Press, New York and London.
18. COLE K.D., (1962), Joule heating of the upper atmosphere, Aust. J. Phys., 15, 223-235.
19. COLE K.D., (1971), Electrodynamic heating and movement of the thermosphere, Planet. Space Sci. 19, 59-75.
20. REES D., (1971), Ionospheric winds in the auroral zone, J. Brit. Interplan. Soc. 24, 233-346.
21. REES D., (1973), Neutral wind structure in the thermosphere during quiet and disturbed

geomagnetic period, in Physics and Chemistry of Upper Atmospheres (Edited by B.M. McCormac), 11-23, Reidel, Dortrecht.
22. PEREIRA E., M.C. KELLEY, D. REES, I.S. MIKKELSON, T.S. JORGENSEN and T.J. FULLER-ROWELL, (1980), Observations of neutral wind profiles between 115 and 176 km altitude in the dayside auroral oval. J. Geophys. Res. 85, 2935-2940.
23. HAKEL M., R.A. WOLF, P.H. REIFF, R.W. SPIRO, W.J. BURKE, F.J. RICH and M. SMIDDY, (1981), Quantitative Simulation of a Magnetospheric Substorm, 1, Model Logic and Overview, J. Geophys. Res. 86, 2217-2241.
24. FULLER-ROWELL T.J., D. REES, S. QUEGAN, R.J. MOFFETT, and G.J. BAILEY, (1987), The thermospheric response and feedback to magnetospheric forcing. Extended Abstract, Symposium on Quantitative Modeling of Magnetosphere-Ionosphere coupling processes. Convenors: Y. Kamide and R.A. Wolf, March 9-13, 1987 Kyoto Sangyo University. p20.
25. CHIU Y.T., (1975), An improved phenomenological model of ionospheric density, J. Atmos. Terr. Phys. 37, 1563-1570.
26. QUEGAN S., G.J. BAILEY, R.J. MOFFETT, R.A. HEELIS, T.J. FULLER-ROWELL, D. REES and R.W. SPIRO, (1982), Theoretical study of the distribution of ionization in the high-latitude ionosphere and the plasmasphere : First results on the mid-latitude trough and the light-ion trough, J. Atmos. Terr. Phys. 44, 619-640.
27. FULLER-ROWELL, T.J., D. REES, S. QUEGAN, G.J. BAILEY AND R.J. MOFFETT, (1984), The effect of realistic conductivities on the high-latitude thermospheric circulation, Planet. Space Sci. 32, 469-480.
28. FULLER-ROWELL T.J., S. QUEGAN, D. REES, R.J. MOFFETT, G.J. BAILEY, (1987b), Interactions between neutral thermospheric composition and the polar ionosphere using a coupled ionosphere-thermosphere model, J. Geophys. Res. 92, 7744-7748.
29. FULLER-ROWELL T J., D. REES, S. QUEGAN, R.J. MOFFETT, G.J. BAILEY, (1988), Simulations of the seasonal and universal time variations of the thermosphere and ionosphere using a coupled, three-dimensional, global model. PAGEOPHYS, 127, 189-217.
30. REES D. and T.J. FULLER-ROWELL (1987), Comparison of theoretical models and observations of the thermosphere and ionosphere during extremely disturbed geomagnetic conditions during the last solar cycle. Adv. Space Res. 7, 827-838, Pergamon.
31. REES D. and T.J. FULLER-ROWELL (1987), A Theoretical Thermosphere Model for CIRA, Adv. Space Res. 7, 10185-10197, Pergamon.
32. SOJKA J.J. and R.W. SCHUNK, (1983), A theoretical study of the high-latitude F-region response to magnetospheric storm inputs, J. Geophys. Res. 88, 2112-2122.
33. QUEGAN S., G.J. BAILEY, R.J. MOFFETT and L.C. WILKINSON, (1986), Universal time effects on the plasma convection in the geomagnetic frame, J. Atmos. Terr. Phys. 48, 25-40.
34. WATKINS B.J., (1978), A numerical computer investigation of the polar F-region, Planet. Space Sci., 26, 559-569.
35. ALLEN B.T., G.J. BAILEY and R.J. MOFFETT, (1986), Ion distributions in the high-latitude topside ionosphere, Ann. Geophysicae 4 A, 97-106.
36. MCCORMAC F.G. and R.W. SMITH, (1984), The influence of the Interplanetary Magnetic Field Y component on ion and neutral motions in the polar thermosphere, Geophys. Res. Lett., 11, 935-938.
37. *FULLER-ROWELL T.J. and D. REES, (1984), Interpretation of an anticipated long-lived* vortex in the lower thermosphere following simulation of an isolated substorm. Planet. Space Sci. 32, 69-85, 1984.
38. HEDIN A.E., (1983), A revised thermospheric model based on mass spectrometer and incoherent scatter data, J. Geophys.Res. 88, 10170-10188.
39. HEDIN A.E. (1988), CIRA 1986, Chapter 1, The Empirical Model Atmosphere: Atmospheric Model in the Region 90 to 2000 km, pub. Pergamon.
40. KILLEEN T.L., B. NARDI, F.G. MCCORMAC, J.W. MERIWETHER Jnr., J.P THAYER., R.G. ROBLE T.J. FULLER-ROWELL and D. REES, (1989), Lower Thermospheric Structure and Dynamics Inferred from Satellite and Ground-based Fabry-Perot Observations of the O(1S) Green line Emission. Adv. Space Res. (in press).
41. NYGREN, T., L. JALONEN, J. OSKMAN and T. TURINEN, The role of electric field and neutral wind direction in the formation of sporadic-E layers, (1984) J. Atmos. Terr. Phys., 46, 373.
42. REES D. T.J. FULLER-ROWELL, M.F. SMITH, R. GORDON, T.L. KILLEEN, P.B. HAYS, N.W. SPENCER, L.E. WHARTON and N.C. MAYNARD, (1985), The Westward Thermospheric Jet Stream of the evening auroral oval. Planet Space Sci., 33, 425-456.

Figure 3 is shown in both colour and black and white to show the grey-scale rendering of the colours in the remainder of the figures.

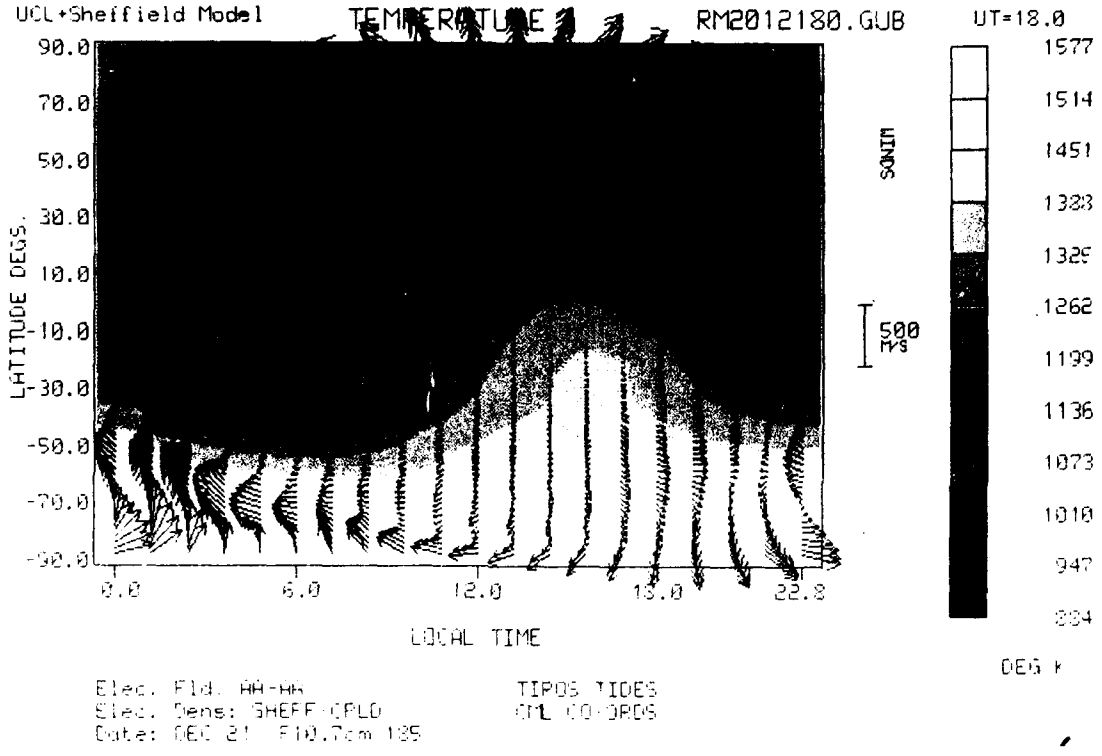



Figure 1. Global distributions of upper thermospheric temperature and wind velocity, pressure level 12, Computed for the June solstice, by the Coupled Ionosphere-thermosphere Model, using the AA convection field, High solar activity, $F_{10.7}$ = 185; Moderate geomagnetic activity, Kp = 3; lower atmosphere propagating semi-diurnal tides.

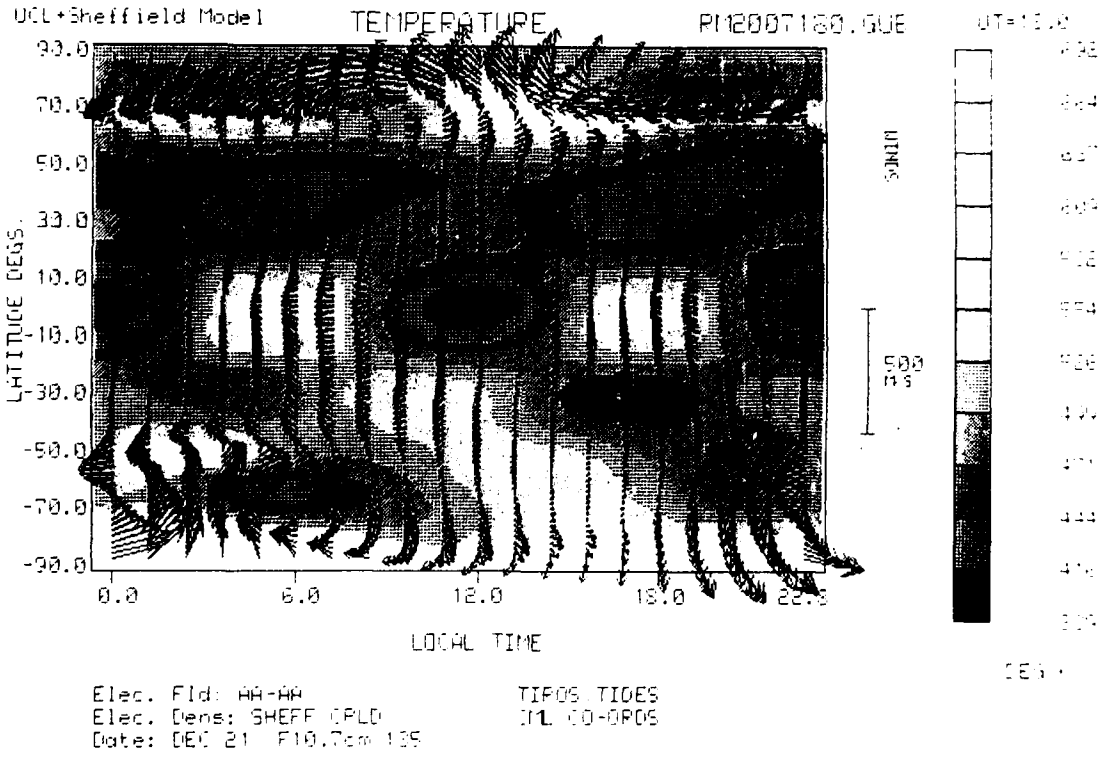



**Figure 2.** Same as Figure 1, showing the global E-Region (125km) response.

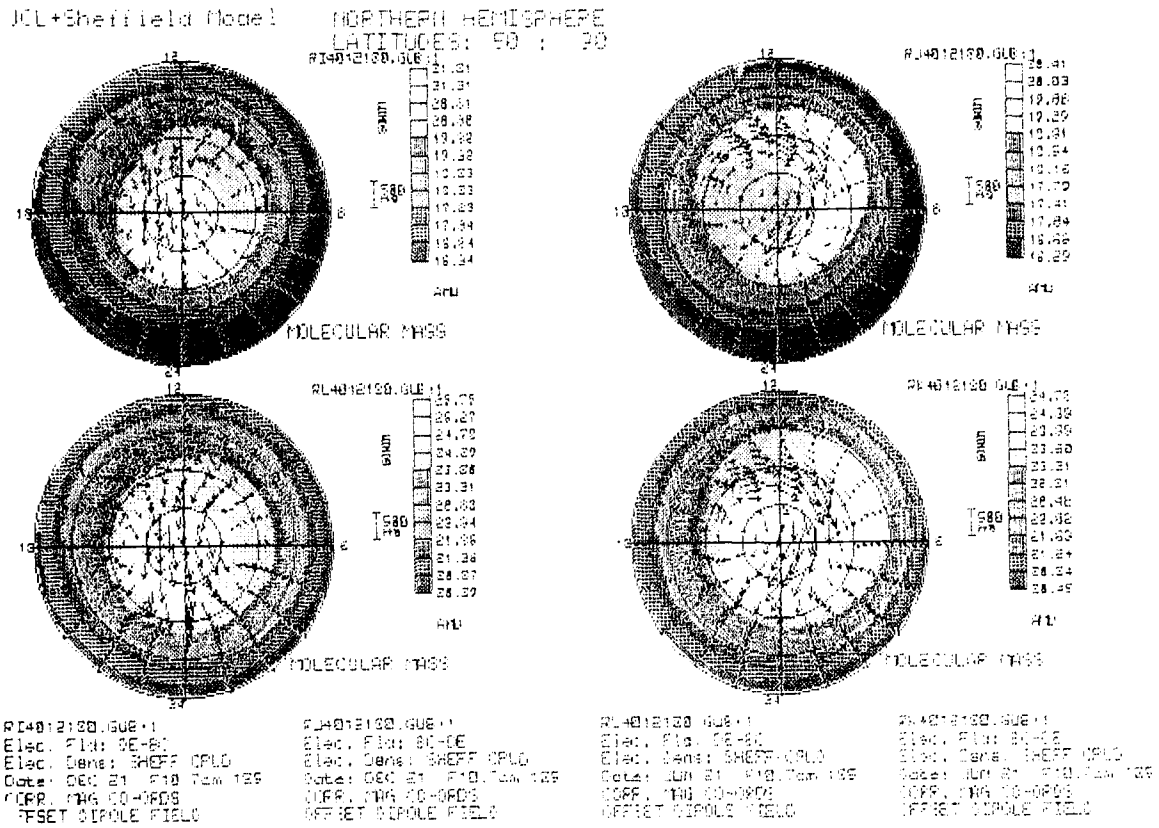



**Figure 3. A comparison, at F-Region altitudes (Pr Lv 12), of neutral composition and winds in the North polar region produced in four simulations.**

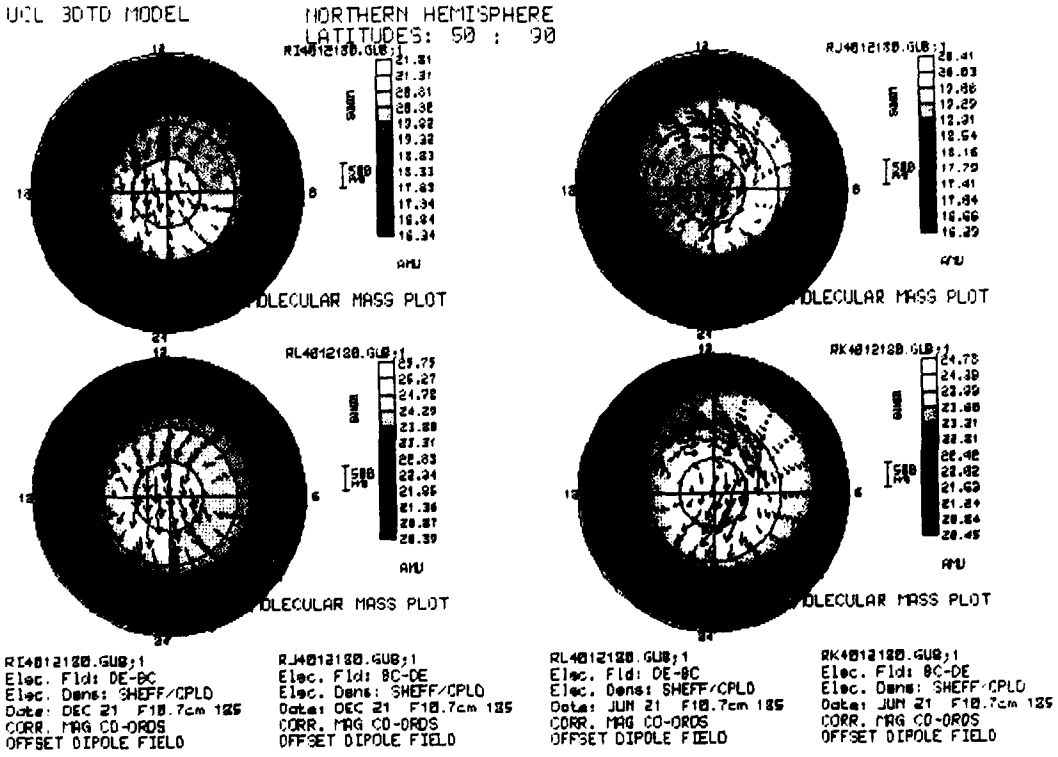



Figure 3. A comparison, at F-Region altitudes (Pr Lv 12), of neutral composition and winds in the North polar region produced in four simulations. The simulations use the same solar and geomagnetic activity conditions as in Figure 1, but the DE/BC convection fields are used to simulate extreme asymmetric values of the IMF BY component, and also contrast northern summer and winter. The simulations include lower atmosphere semi-diurnal tides. This Figure, and the following Figures have the same format: IMF BY negative on the left hand side, IMF BY positive on the Right Hand Side; Northern Winter is above, and Northern Summer is below. Minimum mean molecular mass values are 16.3 in winter and 20.4 amu in summer; maximum values exceed 21 amu in winter and 25 amu in the summer polar cap.

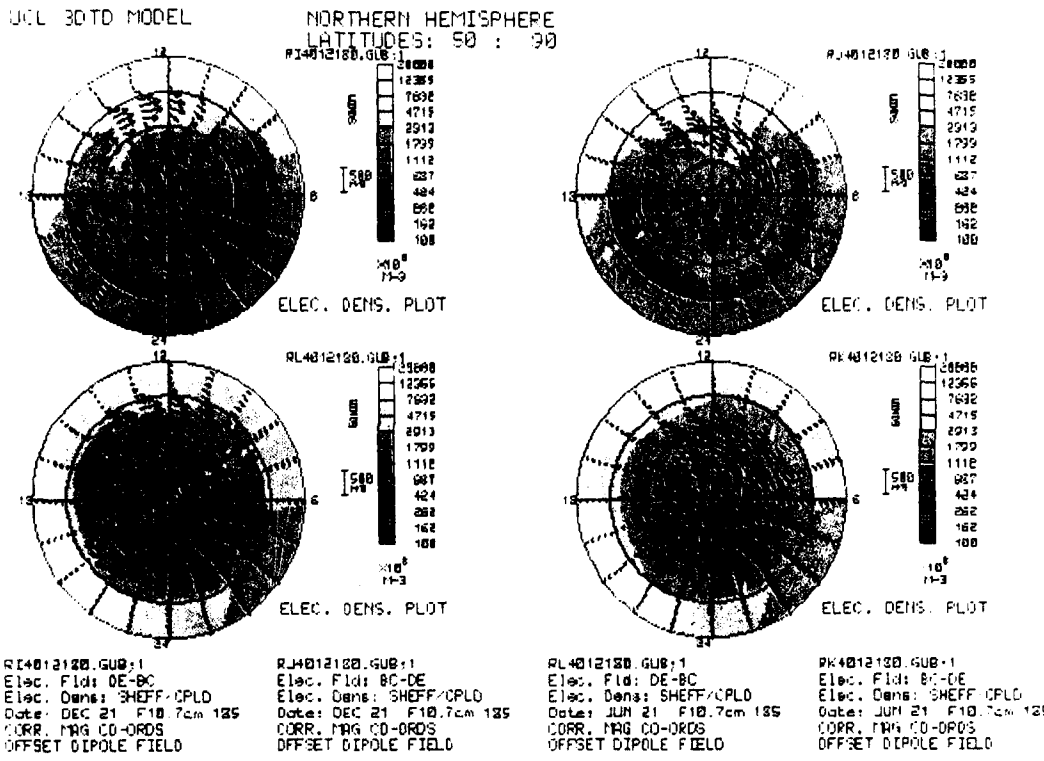



Figure 4. A comparison, for the same four sets of seasonal and IMF BY component conditions as shown in Figure 3, of the corresponding plasma density distributions at F-region altitudes (301 km). There is a contrast between plumes, carried by convection, in the winter simulations (upper figures), and polar 'holes', corresponding to very high values of neutral mean molecular mass, for the summer simulations (lower figures). A common logarithmic scale is used for all four plots, from $1 * 10^{10}$ $m^{-3}$ to $2 * 10^{12}$ $m^{-3}$.

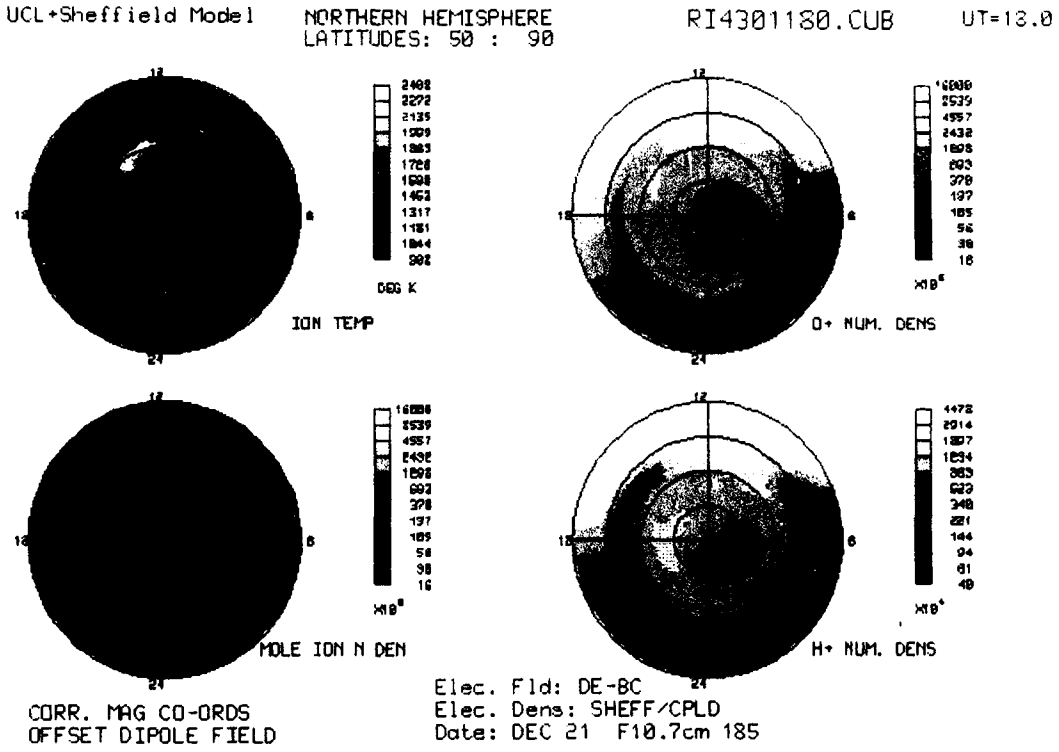



Figure 5. The four sections show the North Polar distributions of Ion Temperature, $O^+$, Molecular Ions and $H^+$ densities at 301 km, for the December simulation, IMF BY -ve, (otherwise as Figure 1). The ion temperature scale is linear 908 to 2408K; $O^+$ and molecular ions use a common logarithmic scale from $1.6 * 10^9$ to $1.6 * 10^{12}$ $m^{-3}$; H+ uses a logarithmic scale from $4 * 10^5$ to $4.47 * 10^7$ $m^{-3}$.

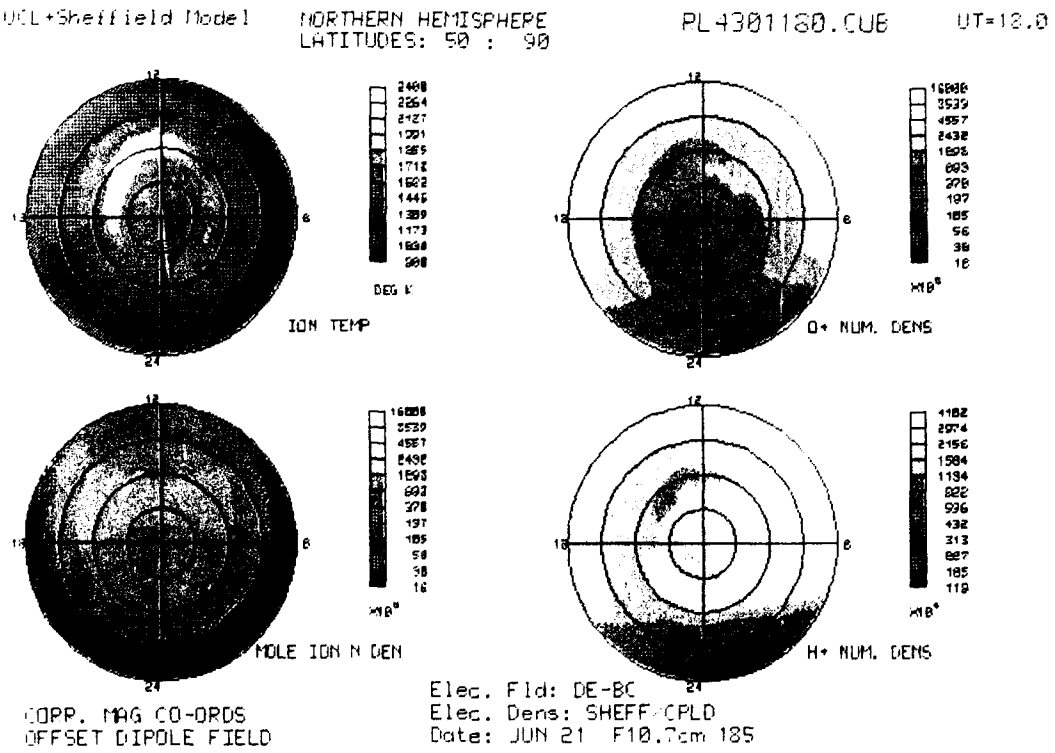



Figure 6. The four sections show the North Polar distributions of Ion Temperature, $O^+$, Molecular Ions and $H^+$ densities at 301 km, for the June simulation, IMF BY -ve, (otherwise as Figure 1). The scales are the same as Figure 5, except that H+ values run from 1.19 * $10^6$ to 4.1 * $10^7$ $m^{-3}$.

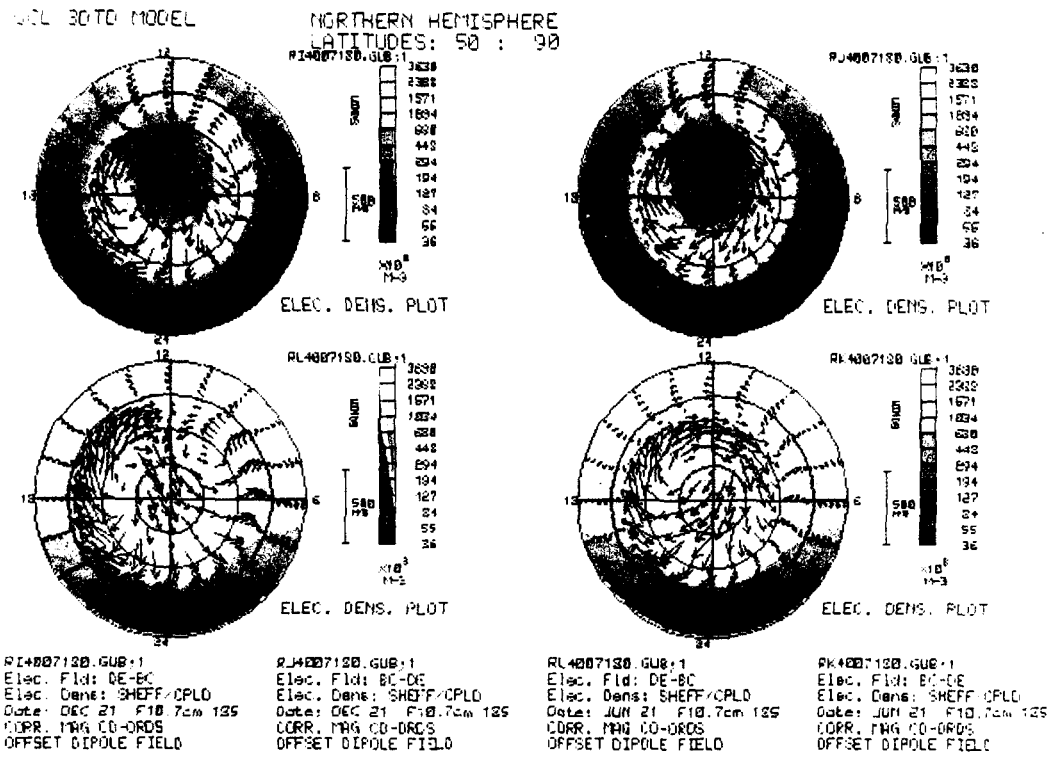



**Figure 7.** A comparison, for the same four sets of seasonal and IMF BY component conditions as shown in Figure 3, of the corresponding plasma density and neutral wind distributions at E-region (pressure level 7, 125 km altitude). The overall and peak plasma density values are higher throughout the summer polar regions, given the combination of solar photoionisation and energetic electron ionisation. The wind distributions show modest response to seasonal changes and marked responses to the sense of IMF BY, while the plasma density distributions show relatively little response to the IMF BY sense (since the precipitation code has no response to IMF BY). A common logarithmic scale is used from 3.6 * $10^{9}$ $m^{-3}$ to 3.6 * $10^{11}$ $m^{-3}$.

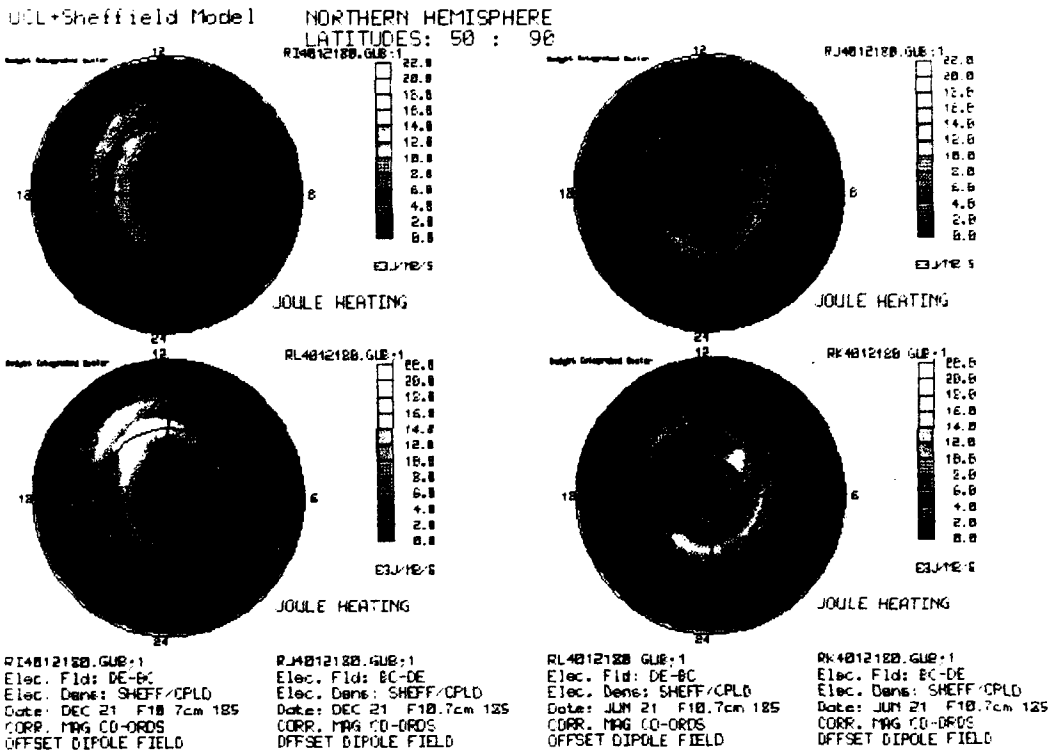



Figure 8. The Height-Integrated Joule Heating rates are shown for the four cases described in Figures 1, 3 and 7. These calculations include the effect of the induced neutral wind dynamo. Summertime values are generally considerably higher, and there is a marked changes in the patterns as the IMF BY component changes sense. The dawn auroral oval/polar cap regions are most strongly heated when IMF BY is positive, while the dusk auroral oval / polar cap, particularly on the dayside, is most strongly heated when IMF BY is negative. A linear scale is used from 0 to 22 mW $m^{-2}$.

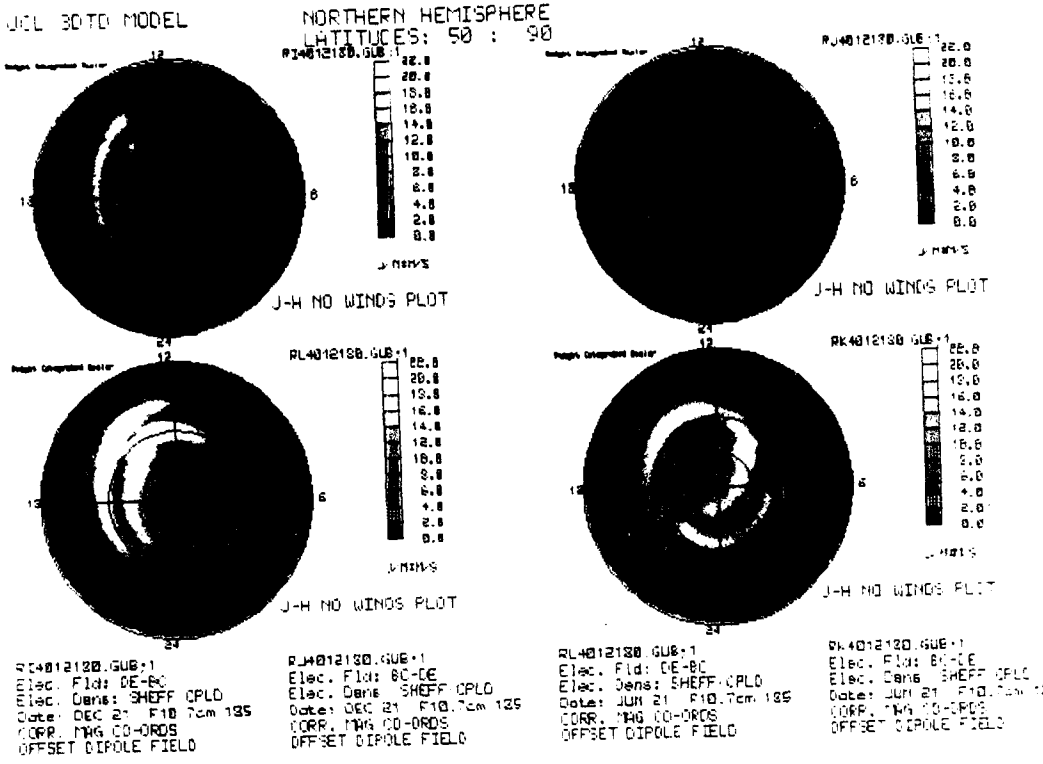



**Figure 9.** Same as Figure 8, with the exception that the induced neutral wind dynamo has been dropped from the calculation. While the overall patterns do not change dramatically, the peak values increase by a factor of 50 %, and in several locations, the emphasis of the highest values shifts from polar cap to auroral oval, or the converse. The scales are as for Figure 8.

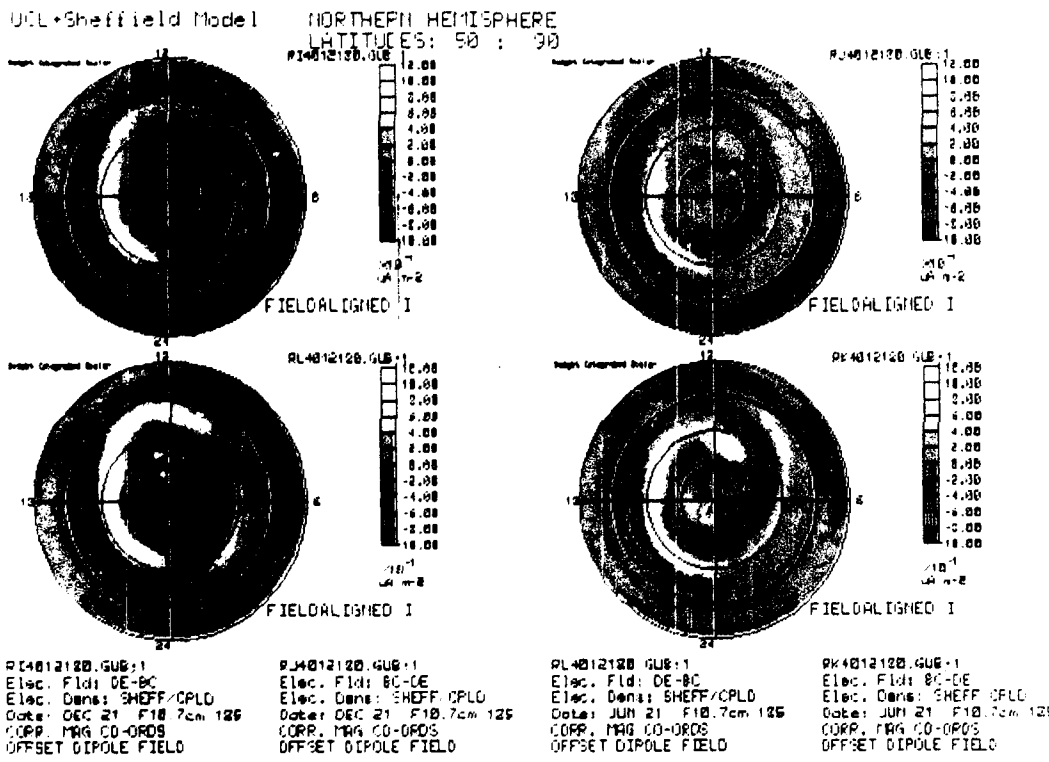



**Figure 10.** For the same simulations as shown in Figures 1, 3 and 7, the field-aligned currents are displayed, calculated from horizontal current convergence or divergence. The major changes occur in the vicinity of the polar cusp, as a result of dependence on the sense of IMF BY, however, the dawn auroral oval and the region of the Harang discontinuity also show considerable BY-dependent changes. A linear scale from -1.0 $\mu A\ m^{-2}$ (downward current to 1.2 $\mu A\ m^{-2}$ (upward current).

## DISCUSSION

**E.Szuszczewicz**

A measure of the merit of a theoretical or numerical model is not only the integrity of its empirical input modules (e.g., convection fields and particle precipitation patterns), but the integrity of its theoretical output in its fit to available data. I worry about the 90-150 km region. For example, MSIS is based on data above 150 km and below 90 km, and simply extrapolates the results of those regions to cover the 90-150 km domain. A thermospheric model therefore has little-to-no data in this region to measure the integrity to specify densities, composition and winds. There is a similar problem with ionospheric density distributions in this region, particularly at night and under conditions of sporadic E. With the 90-150 km region being so critical to things like Pedersen conductivities, current systems, and wind-driven dynamo fields, I wonder what the future holds for global codes, which attempt a self-consistent representation of the magnetosphere-ionosphere-thermosphere system but fare a virtually non-existent data base in the ever-critical 90-150 km domain. Would you comment on this problem?

**Author's Reply**

The region below 250 km is a serious problem because it is a hard place to make measurements. Rockets can probe that region, but it would be very expensive to build an empirical data base that way. The Upper Atmospheric Research Satellite will not really address the 90-150 km region very seriously because almost nothing gives a good optical signature in this region. The CEDAR program in the United States will help, but there will not be a major improvement in the near future in comparison to the data that are needed.

# STORMTIME COUPLING OF THE NEUTRAL AND IONIZED UPPER ATMOSPHERE AT MIDDLE LATITUDES

G.W. Prölss

Institut für Astrophysik
und Extraterrestrische Forschung
Universität Bonn
Auf dem Hügel 71
D-5300 Bonn 1, Fed. Rep. Germany

SUMMARY

Our present knowledge about the interaction between magnetic storm associated changes in the neutral gas composition and ionospheric storms is reviewed. After a brief description of ionospheric disturbance effects, it is shown that the negative phase is caused by a decrease of the atomic oxygen to molecular nitrogen (oxygen) density ratio. Local time variations are discussed as an example of systematic changes common to both disturbance phenomena. It is shown that negative ionospheric storms may be confused with ionospheric trough effects which are not produced by neutral composition changes. There is also no association between composition perturbations and positive ionospheric storm effects which again are attributed to the transport of ionization.

## 1. INTRODUCTION

Magnetic storm associated perturbations of the neutral and ionized upper atmosphere remain a challenging topic of solar-terrestrial physics. They not only have a profound influence on the global morphology of the upper atmosphere but may also severely degrade subionospheric radio communication and satellite ephemeris predictions. This review is concerned with the coupling of neutral atmospheric and ionospheric disturbance phenomena. Specifically, the interaction between storm-associated changes in the neutral gas composition and ionospheric storm effects is investigated. The material is thereby organized in the following way: Section 2 briefly reviews some basic properties of ionospheric storms. The cause-effect relationship between neutral composition changes and negative ionospheric storms is demonstrated in Section 3. A discussion of ionospheric trough effects and their relation to negative ionospheric storms (Section 4) follows. Finally, Section 5 briefly reviews the generation of positive ionospheric storms.

## 2. BASIC MORPHOLOGY OF IONOSPHERIC STORMS

Early studies established the fact that at middle latitudes, the maximum electron density may be enhanced (positive ionospheric storm effects) and/or depressed (negative ionospheric storm effects) during a magnetic storm. Figure 1 illustrates both disturbance effects by showing the time variation of the maximum electron density as observed at two mid-latitude stations during a magnetic storm. The monthly median serves as a quiet-time reference. As is evident, the ionosphere above Brisbane exhibits a large reduction of the maximum electron density, whereas the ionosphere above Pt. Arguello shows a considerable enhancement of ionization. Since measurements of the total electron content of the ionosphere show very similar variations, the observed changes in the maximum electron density reflect the behavior of the entire F-layer and are not simply due to a distortion of the ionization height profile.

The completely different disturbance behavior of the ionosphere above two comparable mid-latitude locations indicates the complex morphology of ionospheric storms. Thus strong longitudinal and latitudinal asymmetries exist. In addition, storm patterns may vary considerably from one event to the next. Apart from this apparently irregular behavior, certain systematic variations are also observed. For example, the north-south asymmetry illustrated in Fig. 1 may be attributed to seasonal variations, with negative storm effects extending to much lower latitudes in the summer hemisphere. Other regular variations include changes with magnetic storm intensity, geomagnetic position, and local time (e.g. [1], and references therein).

## 3. COMPOSITION CHANGES AND NEGATIVE STORM EFFECTS

More than 30 years ago, it was suggested that the negative phase of ionospheric storms is caused by changes in the neutral gas composition which were assumed to be an integral part of atmospheric storms [2]. This explanation has been fully supported by satellite-borne mass spectrometer measurements (e.g. [3]). Figure 2 shows the basic latitudinal structure of a large thermospheric storm. Magnetic storm induced changes in the exospheric temperature, in the argon, molecular nitrogen, atomic oxygen, and helium densities, in the molecular nitrogen to atomic oxygen density ratio, and in the total mass density are plotted as a function of magnetic latitude. The departures from quiet-time conditions are presented in the form of relative changes. Thus R(n) is defined as the stormtime value of a quantity n divided by the corresponding quiet-time value, and R(n) = 1 serves as reference, meaning no change with respect to quiet times.

The prominent feature of Fig. 2 is the well-developed composition disturbance zone at high and middle latitudes. It is marked by a large temperature increase, a large increase in the argon and molecular nitrogen densities, a moderate but still significant decrease in the atomic oxygen density, and a large decrease of the helium density. With respect to ionospheric perturbations, it is the decrease of atomic oxygen and the concurrent increase of the molecular nitrogen density which are of immediate interest. This is because a decrease of atomic oxygen will cause a decrease of the ionization production rate, and an increase in the molecular species will cause an increase in the ionization loss rate. Thus

both density changes combine to produce a decrease of the F-layer ionization density. The molecular nitrogen to atomic oxygen density ratio is therefore a suitable parameter to study the coupling between both disturbance phenomena.

One would expect that whenever an ionosonde station is located inside a region of disturbed $N_2/O$ composition, it will experience negative ionospheric storm effects, and this is indeed observed. Figure 3 shows a comparison between the latitudinal distribution of neutral atmospheric and ionospheric disturbance effects. Relative changes in the $N_2/O$ density ratio indicate the latitudinal structure of the atmospheric disturbance zone. The ionospheric response to the composition disturbance is illustrated in the lower part of the figure using critical frequency data from three ionosonde stations located along the satellite pass. As is evident, both stations located inside the atmospheric perturbation zone show significant negative storm effects, whereas the station located outside the perturbation zone does not observe any depression in the ionization density. This kind of comparison clearly documents the close correlation between composition changes and negative ionospheric storm effects.

To establish the cause-effect relationship between both phenomena also in a quantitative way, a numerical simulation of negative storms was performed using observed composition data as input. The results of this calculation are illustrated in Fig. 4. The upper part shows relative day-to-day variations in the $O/N_2$ density ratio as observed near an ionosonde station during a 6-day time interval. These composition data were introduced into an ionospheric model to reproduce the day-to-day variation in the maximum electron density observed at this station. The procedure adopted was to adjust the exospheric temperature and the atomic oxygen density at the lower boundary level of the atmospheric model in such a way that at the time of the satellite pass (about 0900 LT), the model values for O and $N_2$ agreed with the measured ones. For simplicity it was assumed that, superimposed on the daily variation, there is a smooth transition from one adjusted value to the next. The results of this computation are presented in the lower part of Fig. 4. Also shown for comparison is the observed variation of the maximum electron density. The good agreement between theoretical prediction and actual measurement is evident.

Additional evidence for the close coupling between neutral composition changes and negative storm effects is derived from the fact that both phenomena exhibit the same systematic variations. These include systematic changes with the disturbance intensity, with magnetic position, with local time, and with season (e.g. [11]). There also may be a dependence on solar activity and on the By component of the interplanetary magnetic field, but these changes remain to be firmly established. In what follows, only the local time variation will be discussed.

One of the more prominent characteristics of negative ionospheric storms is that they most frequently commence in the nighttime - early morning sector but rarely during the daytime. This is illustrated in Fig. 5, which shows the local time distribution of onset times of negative ionospheric storms as observed at a number of mid-latitude stations in the course of a year. Although these onset times were determined by inspection and therefore involve some subjective judgment, and although some of the data points come from adjacent stations, it is safe to conclude that most storms are first observed between, say, 4 and 8 in the morning, but rarely commence between 11 and 17 hours local time. This local time asymmetry is consistent with the idea that the major transfer of disturbance energy from the polar to the mid-latitude region occurs in the early morning sector. This point is illustrated schematically in Fig. 6. At storm onset, the strong expansion of the atmospheric disturbance toward lower latitudes is restricted to the morning sector. It is only later in the post-expansion phase that the residuals of the morning disturbance are also seen in the local afternoon sector. This implies that the disturbance magnitude in the morning sector should be significantly larger than that in the afternoon/evening sector, and this is indeed observed (e.g. [7] and [8]).

At least two aspects of the above model need further clarification. First, this model rests on the assumption that the disturbance effects are transported from high to middle latitudes and are not produced locally. Although consistent with present experimental evidence, this assumption cannot yet be considered as proven. Secondly, the disturbance transport itself requires further study, and different mechanisms - like large-scale circulation, large-scale convection, waves and cyclone-type wind systems - have been discussed in this connection.

## 4. NEGATIVE STORM EFFECTS AND IONOSPHERIC TROUGHS

Since both negative ionospheric storms and ionospheric troughs are marked by a decrease in the ionization density, and since both phenomena move toward lower latitudes during disturbed conditions, their signatures may be confused. This is especially true for higher mid-latitude locations where both effects may be observed. Figure 7 illustrates this point using critical frequency data obtained at Ottawa ($58^0$ invariant latitude) during a disturbed day. Following an intermediate increase (positive storm effect) of the electron density in the afternoon, there is a sharp drop of the density in the evening, suggesting the onset of a negative ionospheric storm. This, however, is not the case. Satellite data obtained during this event instead suggest that Ottawa had rotated into the trough region and that the trough had moved and expanded beyond the latitude of Ottawa, causing the abrupt drop in the [illegible] density observed. Similar events were documented before and were interpreted in the same way (e.g. [9]).

The distinction between negative ionospheric storms and trough effects implies the assumption that both phenomena have different origins. According to the current literature (e.g. [10 - 13] and references therein), this is indeed the case. Thus neutral gas composition changes are generally considered as unimportant for the trough formation. However, actual evidence in support of this claim is scarce [14-15]. Therefore an additional study was undertaken to verify this point. The results may be summarized as follows: (1) In the winter evening sector and during moderately active conditions, there usually exists a well-developed ionospheric trough at subauroral latitudes which features a steep drop in the electron density at the equatorward side and a less steep rise at the poleward side. (2) There also exists as a permanent feature in this region a localized composition disturbance zone which is marked by a moderate increase in the molecular nitrogen density. (3) A comparison of the latitudinal distribution shows that the composition disturbance zone is usually located poleward of the trough. Also, there is no similarity between the ramp-like increase of the $N_2$ density and the v-shaped structure of the trough.

(4) Whereas the electron density shows a large decrease (≈ factor 10), the $N_2$ density is only moderately enhanced (≈ factor 2). This data comparison convincingly demonstrates that ionospheric troughs are not caused by changes in the neutral gas composition. This also applies for the nighttime/early morning sector where the trough is less deep and the $N_2$ density increase is somewhat larger.

## 5. POSITIVE IONOSPHERIC STORMS

It has been repeatedly suggested that positive ionospheric storms are also caused by changes in the neutral gas composition (e.g. [16-18]). This contention, however, is not supported by the data. Figure 8 shows another example of the distribution of ionospheric storm effects in relation to the latitudinal structure of an atmospheric disturbance. It is evident that the positive storm effects observed at two of the stations (Boulder and Winnipeg) are not associated with any significant changes in the neutral gas composition.

Further evidence for this lack of correlation is presented in Fig. 9. Day-to-day variations of the $O/N_2$ density ratio are compared to changes in the maximum electron density. No correlation is found between the smoothly varying composition and the positive storm effects observed between October 28 and 30.

These data strongly suggest that positive storms are caused by transport of ionization. It was demonstrated that both meridional winds and zonal electric fields can lift ionization into regions of reduced loss rates, thereby causing an increase in the density (e.g. [19]). However, positive storm effects were also modeled using horizontal electrodynamic drifts [20], and, during nighttime, using downward transport of ionization [21]. Since all mechanisms can produce positive storm effects, it is quite possible that different mechanisms are operative at different times or even simultaneously.

Figure 10 illustrates a special kind of positive storm effect, namely that of short duration sometimes observed at middle latitudes in association with substorm activity. In this particular case, the positive storm effect is most probably caused by an upward transport of ionization. Figure 11 shows that in response to the substorm activity, the ionospheric F-layer is lifted upward by tens of kilometers. Subsequently, the marked increase in ionization density is observed. A possible explanation for this increase in layer height are meridional winds carried from the polar region toward lower latitudes by large-scale traveling atmospheric disturbances (e.g. [22-25]).

Accordingly, if it is assumed that positive disturbance effects are caused by winds and/or electric fields, the principle interactions between magnetospheric, neutral atmospheric, and ionospheric storms may be summarized as shown in the block diagram of Fig. 12. To keep the diagram simple, possibly important interactions have been omitted. For example, electric fields will also modify the wind system and vice versa; or a reduction of the ionospheric plasma density may result in larger wind speeds; and at times, positive and negative storm effects may be competing phenomena.

## REFERENCES


1. Prölss, G.W., Magnetic storm associated perturbations of the upper atmosphere: recent results obtained by satellite-borne gas analyzers, Rev. Geophys. Space Phys. 18, 1980, 183-202.
2. Seaton, M.J., A possible explanation of the drop in F-region critical densities accompanying major ionospheric storms, J. atmos. terr. phys. 8, 1956, 122-124.
3. Prölss, G.W., and U. von Zahn, Esro 4 gas analyzer results 2. Direct measurements of changes in the neutral composition during an ionospheric storm, J. Geophys. Res. 79, 1974, 2535-2539.
4. Prölss, G.W., and M. Roemer, Thermospheric storms, Adv. Space Res. 7, No. 10, 1987, 223-235.
5. Jung, M.J., and G.W. Prölss, Numerical simulation of negative ionospheric storms using observed neutral composition data, J. atmos. terr. Phys. 40, 1978, 1347-1350.
6. Prölss, G.W., and U. von Zahn, On the local time variation of atmospheric-ionospheric disturbances, Space Res. 18, 1978, 159-162.
7. Babcock, R.R., Evans, J.V., Effects of geomagnetic disturbances on neutral winds and temperatures in the thermosphere observed over Millstone Hill, J. Geophys. Res. 84, 1979, 5349-5354.
8. Prölss, G.W., Local time dependence of magnetic storm effects on the atmosphere at middle latitudes, Ann. Geophys. 2, 1984, 481.
9. Buonsanto, M.J., M. Mendillo, and J.A. Klobuchar, The ionosphere at L = 4: average behavior and the response to geomagnetic storms, Ann. Geophys. 35, 1979, 15-26.
10. Schunk, R.W., P.M. Banks, and W.J. Raitt, Effects of electric fields and other processes upon the nighttime high-latitude F-layer, J. Geophys. Res. 81, 1976, 3271-3282.
11. Moffett, R.J. and S. Quegan, The mid-latitude trough in the electron concentration of the ionospheric F-layer: a review of observations and modelling, J. atmos. terr. Phys. 45, 1983, 315-343.
12. Schlegel, K., A case study of a high latitude ionospheric electron density depletion, J. atmos. terr. Phys. 46, 1984, 517-520.
13. Senior, C., J.R. Sharber, O. de la Beujardiere, R.A. Heelis, D.S. Evans, J.D. Winningham, M. Sugiura, and W.R. Hoegy, E and F region study of the evening sector auroral oval: A Chatanika / Dynamics Explorer 2 / NOAA 6 comparison, J. Geophys. Res. 92, 1987, 2477-2494.
14. Raitt, W.J., U. von Zahn, and P. Christophersen, A comparison of thermospheric neutral gas heating and related thermal and energetic plasma phenomena at high latitudes during geomagnetic disturbances, J. Geophys. Res. 80, 1975, 2277-2288.
15. Wrenn, G.L., and W.J. Raitt, In situ observations of mid-latitude ionospheric phenomena associated with the plasmapause, Ann. Geophys. 31, 1975, 17-28.
16. Mayr, H.G., I. Harris, and N.W. Spencer, Some properties of upper atmospheric dynamics, Rev. Geophys. Space Phys. 16, 1978, 539-565.
17. Rishbeth, H., R. Gordon, D. Rees, and T.J. Fuller-Rowell, Modelling of thermospheric composition changes caused by a severe magnetic storm, Planet. Space Sci. 33, 1985, 1283-1301.
18. Danilov, A.D., L.D. Morozova, Tc. Dachev, and I. Kutiev, Positive phase of ionospheric storms and its connection with the dayside cusp, Adv. Space Res. 7, No. 8, 1987, 81-88.

19. Rishbeth, H., and O.K. Garriott, Introduction to ionospheric physics, Academic Press, New York and London, 1969, 151-154.
20. Anderson, D.N., Modeling the midlatitude F-region ionospheric storm using east-west drift and a meridional wind, Planet. Space Sci. 24, 1976, 69-77.
21. Park, C.G., and P.M. Banks, Influences of thermal plasma flow on the mid-latitude nighttime $F_2$ layer: Effects of electric fields and neutral winds inside the plasmasphere, J. Geophys. Res. 79, 1974, 4661-4668.
22. Roble, R.G., A.D. Richmond, W.L. Oliver, and R.M. Harper, Ionospheric effects of the gravity wave launched by the September 18, 1974, sudden commencement, J. Geophys. Res. 83, 1978, 999-1009.
23. Prölss, G.W., and M.J. Jung, Traveling atmospheric disturbances as a possible explanation for daytime positive storm effects of moderate duration at middle latitudes, J. atmos. terr. Phys. 40, 1978, 1351-1354.
24. Tanaka, T., and K. Hirao, Effects of an electric field on the dynamical behavior of the ionospheres and its application to the storm time disturbance of the F-layer, J. atmos. terr. Phys. 35, 1973, 1443-1452.
25. Lanzerotti, L.J., L.L. Cogger, and M. Mendillo, Latitude dependence of ionosphere total electron content: Observations during sudden commencement storms, J. Geophys. Res. 80, 1975, 1287-1306

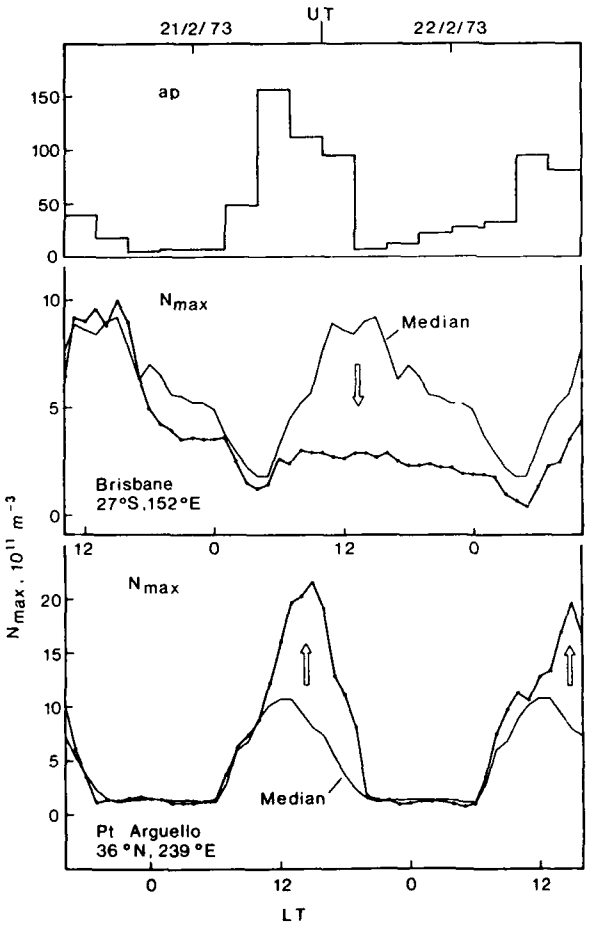



Fig. 1 Magnetic storm induced changes in the ionospheric electron density. Measured F2-layer maximum electron densities ($N_{max}$, heavy dotted lines) obtained at two mid-latitude stations during the magnetic disturbance event of 21/22 February 73 (upper panel) illustrate negative and positive ionospheric storm effects. The monthly median serves as a quiet-time reference (thin lines) [1].

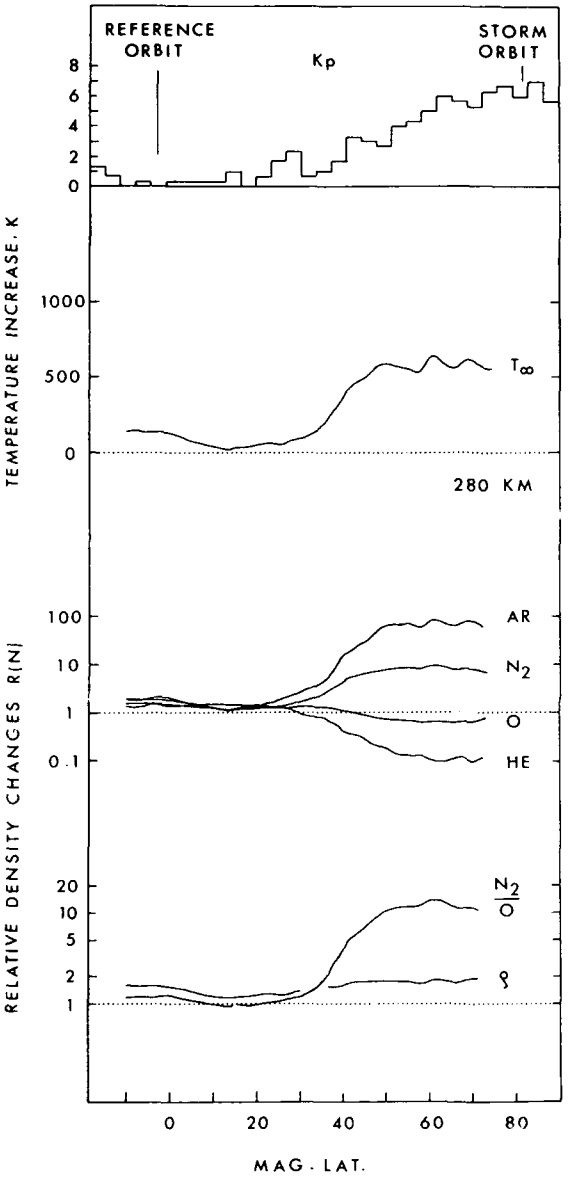



Fig. 2 Basic latitudinal structure of a thermospheric storm. The upper panel shows the development of the magnetic activity during the event. It also indicates the times at which the storm data and the quiet-time reference data were measured. The lower panel presents storm-associated changes in the exospheric temperature ($T_\infty$), in the argon (Ar), molecular nitrogen ($N_2$), atomic oxygen (O), and helium (He) densities, in the molecular nitrogen to atomic oxygen concentration ratio ($N_2/O$), and in the total mass density ($\rho$). The exospheric temperature has been inferred from the nitrogen data. All density data have been adjusted to a common altitude (280 km) and normalized to prestorm conditions. In this kind of presentation, R(n) = 1 serves as a reference, meaning no change with respect to quiet times. The data refer to the morning local time sector (approximately 0900 SLT) and fall conditions. At the equator, the geographic longitude and universal time of observation were 246 $^0$E and 16:52, respectively [4].

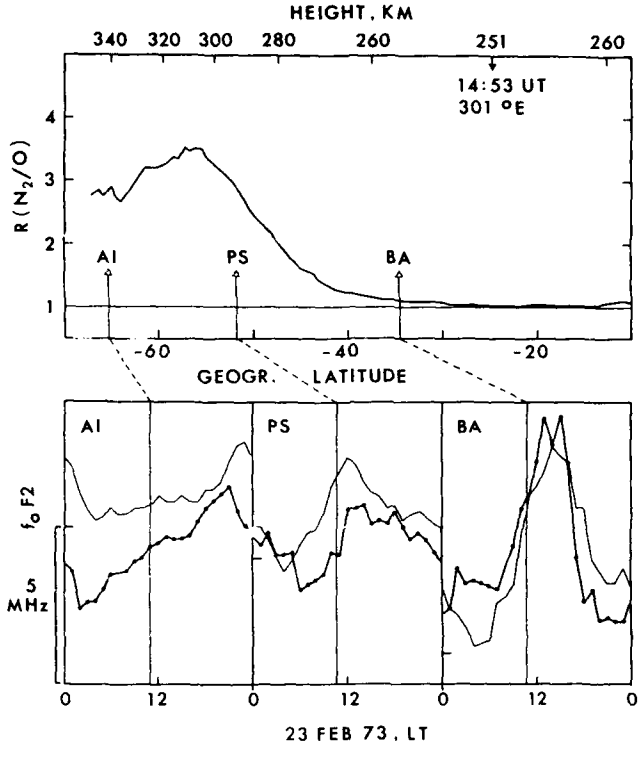



Fig. 3 Comparison between the latitudinal structure of neutral atmospheric and ionospheric disturbance effects. The upper panel shows relative changes in the $N_2/O$ density ratio as observed near 301 °E on 23 February 73. $R(N_2/O) = 1$ serves as the quiet-time reference, and local time of observation is approximately 11 hours. The lower panel shows the local time variation of the F2-layer critical frequency (heavy dotted line) as observed at three ionosonde stations whose relative positions with respect to the atmospheric disturbance zone are indicated by arrows. The monthly median of $f_0F2$ for February 73 (thin line) serves as the quiet-time reference. A tick mark on the ordinate gives the 5 MHz level (for each station separately), and the bar on the left-hand side indicates the scale used for all stations. AI, PS, and BA stand for Argentine Island, Port Stanley, and Buenos Aires [1].

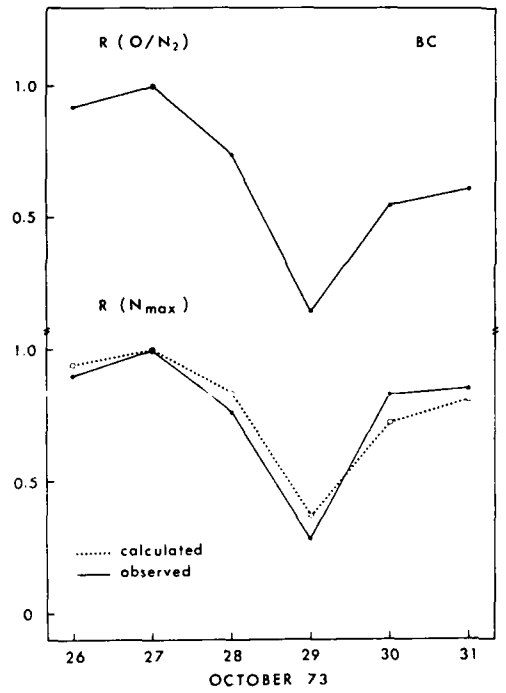



Fig. 4 Comparison between calculated and observed day-to-day variations in the F2-layer maximum electron density. The upper part shows day-to-day variations in the $O/N_2$ density ratio at 260 km altitude as observed near Boulder (BC) during a six-day interval in October 73 (see also Fig. 2). The lower part shows the calculated and observed response of the ionosphere to these composition changes. Variations of the maximum electron density, $N_{max}$, are plotted. Note that all data have been normalized to values obtained on 27 October 73 [5].

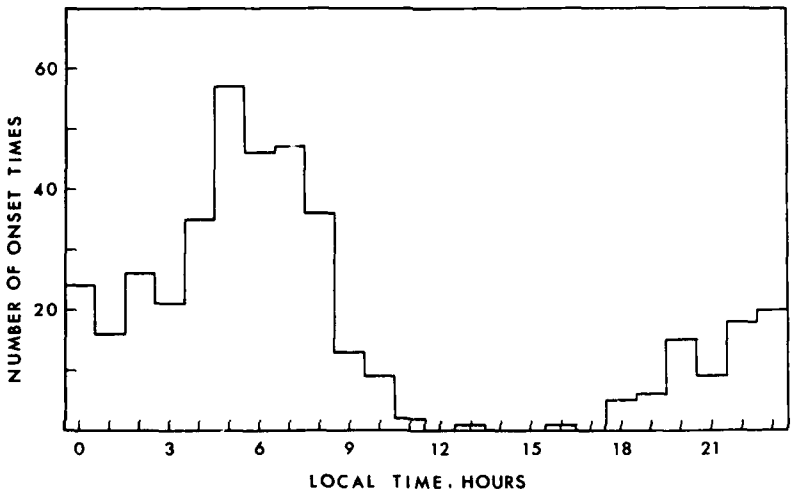



Fig. 5 Local time distribution of onset times of negative ionospheric storms as observed at 24 mid-latitude stations during the year 1973 [6].

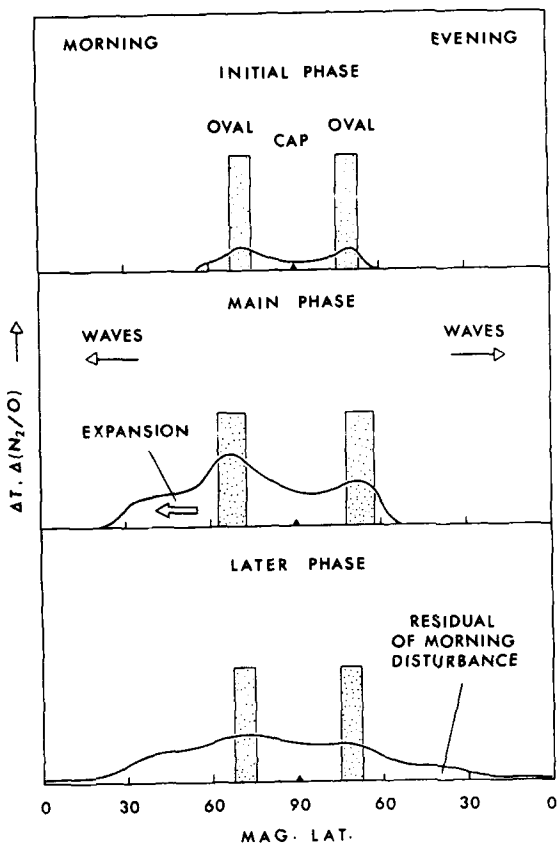



Fig. 6 Possible origin of local time dependent changes of atmospheric storm effects [8].

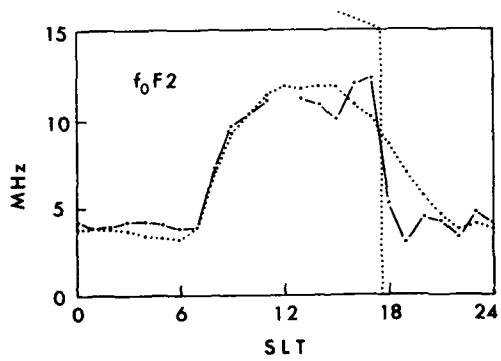



Fig. 7 Trough effects at middle latitudes. Stormtime variation of the critical frequency of the F2 layer (full line) as observed at Ottawa ($58^{0}$ invariant latitude) on 7 December 82. Also shown for comparison is the monthly median (dotted line). The vertical dotted line indicates the time when the DE2 satellite observed a trough near Ottawa.

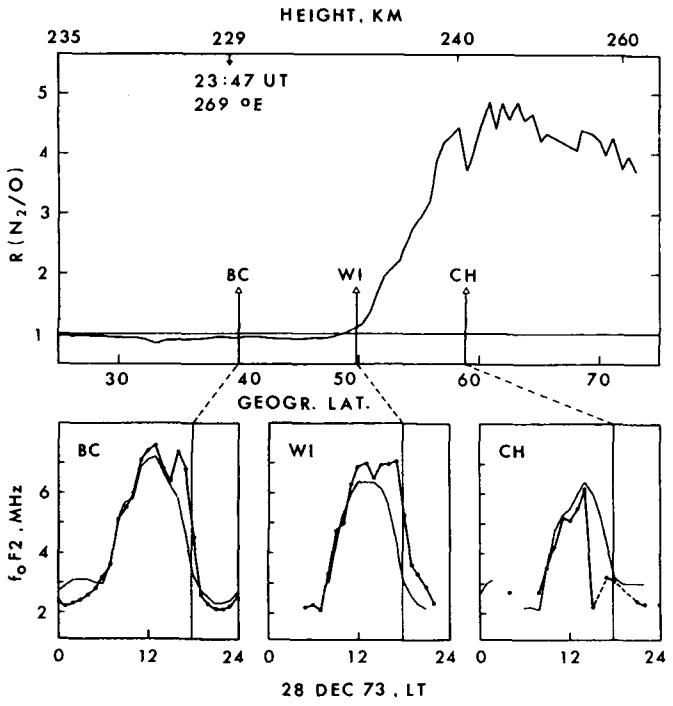



Fig. 8 Comparison between the latitudinal structure of neutral atmospheric and ionospheric disturbance effects. Same as Fig. 3 for data obtained near 269 °E on 28 December 73. Local time of observation is approximately 18 hours [1].

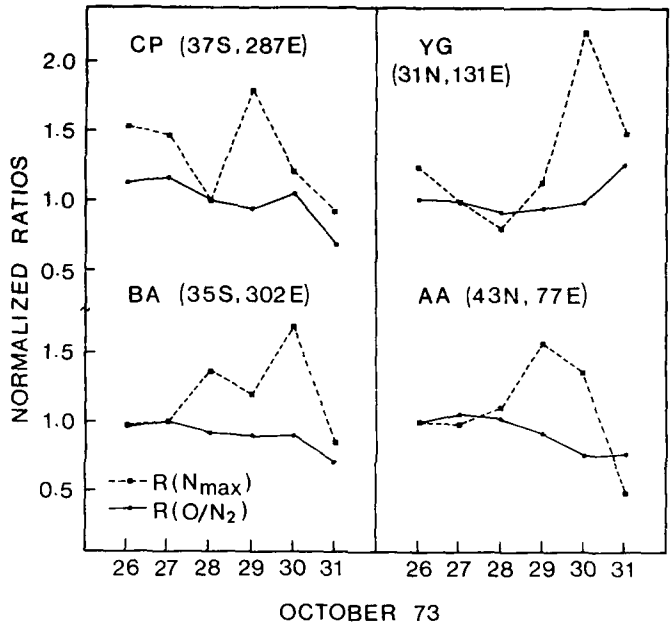



Fig. 9 Neutral composition changes and positive ionospheric storm effects. Day-to-day variations in the F2-layer maximum electron density ($N_{max}$) and in the atomic oxygen to molecular nitrogen ($O/N_2$) density ratio are shown as observed above four mid-latitude stations during a six-day interval in October 73 (see also Figs. 2 and 4). Both data sets have been normalized to values observed on a reference day; in addition, the composition data have been corrected for changes in observation height. Local solar time of observation is approximately 9 hours. Note that in contrast to Fig. 4, all four stations were located outside the neutral composition disturbance zone. CP, BA, YG, and AA stand for Concepcion, Buenos Aires, Yamagawa, and Alma Ata [1].

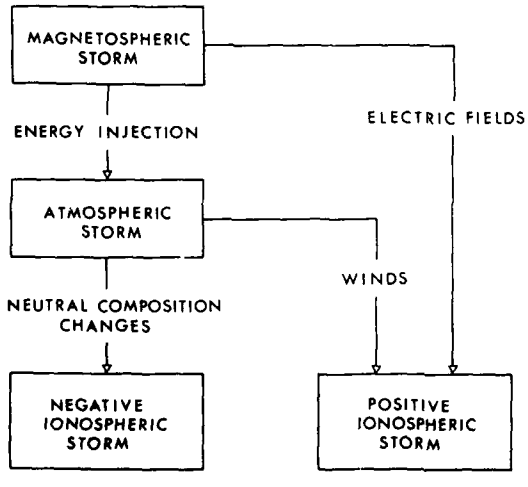



Fig. 12 Causal relation between magnetospheric, neutral-atmospheric and ionospheric disturbance effects [1].

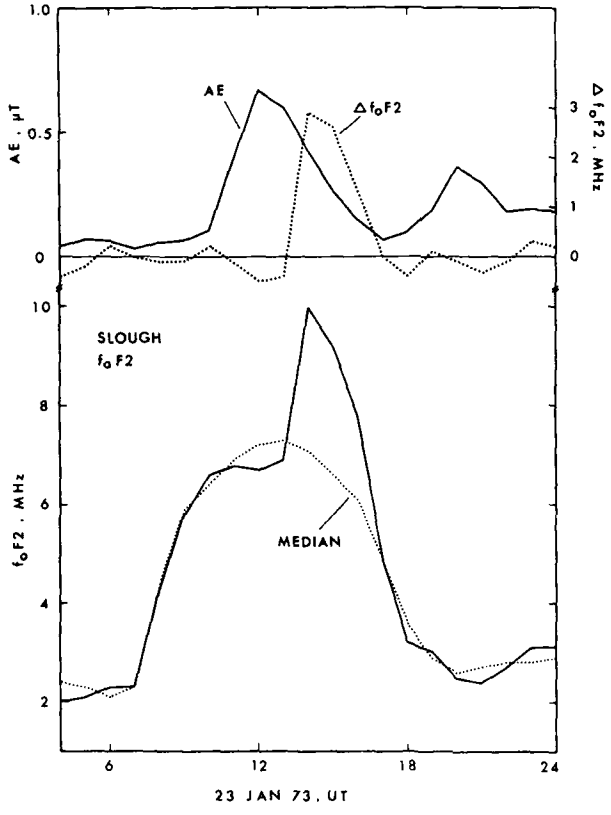



Fig. 10 Example of a short-duration positive ionospheric storm effect. In response to the substorm activity between 10 and 17 UT on 23 January 73 (AE index, upper panel), Slough and other European stations observe a considerable increase in the critical frequency and therefore in the maximum electron density of the F2 layer (lower panel). The median of $f_0F2$ for January 73 serves as the quiet-time reference. To illustrate the time lag between the increase in the AE index and the subsequent increase in $f_0F2$, the change in the critical frequency $\Delta f_0F2 = f_0F2 - (f_0F2)_{median}$ has been superimposed on the variation of the AE index [23].

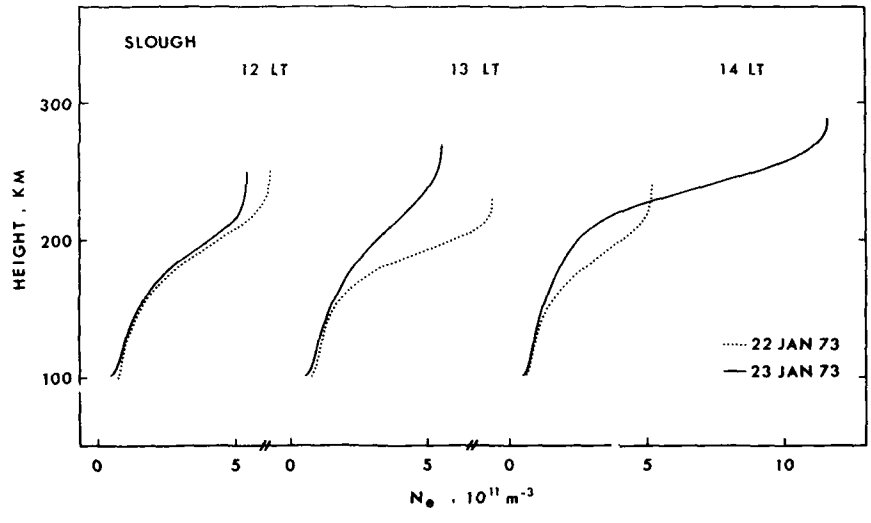



Fig. 11 Changes in the electron density profile during a positive ionospheric storm. The full lines refer to density profiles obtained at Slough during the storm event of 23 January 73 (see also Fig. 10). The dotted profiles serve as a quiet-time reference and were obtained on 22 January 73 [23].

## DISCUSSION

**Y.Tulunay**

I am not commenting that $N_2$ composition changes are not responsible for the trough formation. However, in one of your examples, I believe it was the high-latitude trough not the mid-latitude trough.

**Author's Reply**

Köhnlein and Raitt (*Planet. Space Sci., 25*, 600, 1977) described the position of the mid-latitude trough as a function of magnetic activity and local time. Our study on the trough position based on DE 2 data is in excellent agreement with their findings. We therefore conclude that all results discussed in the present review refer to the mid-latitude trough.

**R.A.Hoffman**

Following a substorm, auroral images from DE-1 have shown a large decrease in the scattered sunlight from midlatitudes in the morning hours. This apparently implies a large decrease in oxygen. Is a cause for this effect known?

**Author's Reply**

To my knowledge, the correlation between reduced oxygen emissions observed by DE 1 and oxygen depletions measured by satellite-borne mass spectrometers at thermospheric altitudes (e.g., DE 2) has not yet been investigated. The fact, however, that both measurement methods reveal atomic oxygen depletions at middle latitudes during magnetospheric activity, which are most pronounced in the early morning sector, indicates to me that both are observing the same phenomenon. If this is the case, a very interesting application of the DE 1 UV data would be to monitor the temporal development and expansion of the thermospheric storm associated oxygen depletion. This cannot be done using data from a low-orbiting satellite.

**D.Rees**

The DE-1 spin scan image shown by Dr. Hoffman shows a large-scale depletion in oxygen 130.4 nm resonance emission. The depletion reflects a major thermospheric disturbance some 6 hours previously. The region of the airglow depletion seems to have a direct connection to negative $F$-region storm disturbance, and neutral gas composition disturbance. It is not clear that the discrete aurora expanded equatorward far enough during this event to force the thermospheric disturbance. The ring-particle decay mechanism (energetic ions) can cause enough disturbance, but the mechanisms are still rather qualitative for inclusion in global models. I think we still have some important problems to investigate and, hopefully, solve!

**Author's Reply**

If the energy from the ring current reaches the thermosphere you should see emission and you don't see the emission.

**D.Anderson**

The author showed a positive storm at Slough. Does this mean that there are just as many positive storms in the European sector as in the American sector?

**Author's Reply**

Although no statistical study has been done, there are apparently more positive storms in the American sector than in the European sector.

**G.Wrenn**

I have a number of questions which I will take up outside the meeting. Rationing myself to just one now, I accept that the trough is a different feature to the negative storm effect, but I have trouble with the idea of a station 'rotating into the trough'. Can you explain what exactly departs from corotation.

**Author's Reply**

It is assumed here that troughs are caused by transport of ionization and that this transport is affected by electric fields. It is also assumed that the pattern of the trough-associated electric fields is quasi-stationary in a solar-geophysical coordinate system. Then the ionosphere above a station will rotate with the station until it comes under the influence of the trough-associated field. This may happen whenever there is a local time asymmetry in the trough position or the trough position moves equatorward. It is in this sense that an ionosonde station (or the ionosphere above it) rotates into the trough region.

**R.Schunk**

I agree with your assessment that $O/N_2$ composition changes play a role in the formation of the main trough. However, several years ago we did a series of ionospheric simulations that showed it is more important to know how the convection and precipitation patterns should be joined in the trough region. Depending on how you matched the boundaries of these patterns, you could get wide or narrow troughs as well as deep or shallow troughs. We found this to be more important than the $O/N_2$ composition effect with regard to trough characteristics.

**Author's Reply**

As stated in this review, neutral composition changes cannot be responsible for the formation of narrow subauroral troughs. However, if it comes to modeling the polar ionosphere, it is important to realize that the convection path followed by an ionospheric volume element may traverse regions of markedly different composition. For example, antisunward convection will take place within the polar composition disturbance zone, but sunward convection may take place outside this region. Unfortunately, presently available neutral atmospheric models are not in a position to reproduce any smaller-scale perturbation effects, including the boundary of the polar atmospheric disturbance zone.

**E.Szuszczewicz**

If you sorted your neutral composition data into dayside and nightside representations, would the storm-time plots as a function of latitude be qualitatively similar but shifted more poleward in the case of the dayside plot?

**Author's Reply**

The latitudinal structure and magnitude of a neutral atmospheric disturbance depend on local time. Significant differences are sometimes observed between, say, the early morning and afternoon perturbation effects (e.g., Prölss, *J. Geophys Res.*, *86*, 2385, 1981).

**H.Rishbeth**

I agree that the positive storm effect is very interesting and complicated. For the "initial phase" positive effects, there may be regional differences: they appear to be especially strong in the North American Sector in the evening. Possible electrodynamic drift and/or influx from the plasmasphere are important. Other such events, e.g., in Europe, seem to be due to equatorward winds, which lift the F2 layer. As regards "main phase" positive effects, modified electric fields could be their cause in low magnetic latitudes. Moreover, our results with the UCL composition model do show an increased atomic/molecular at fixed pressure levels at middle and low latitudes where "downwelling" occurs, and that should have a positive effect on $N_m F_2$. This needs to be quantified.

**Author's Reply**

I do agree with your comment except that I do not consider neutral composition changes an important mechanism for the generation of positive ionospheric storms (see Section 5 of the present review).

# IONOSPHERE-THERMOSPHERE COUPLING AT HIGH LATITUDES: ASPECT ANGLE DEPENDENCE OF NON-THERMAL PLASMAS

by

K.J.Winser, M.Lockwood and A.D.Farmer
Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX, UK

G.O.L.Jones
Physics Dept., University College of Wales, Aberystwyth, Dyfed, UK

D.Rees and A.Aruliah
Dept. of Physics and Astronomy, University College, Gower St., London, UK

## ABSTRACT

Recent observations with the EISCAT incoherent scatter radar have shown large, relatively short lived (~ 1 hour), enhancements in the dayside, auroral convection flows over a side range of latitude. These are larger than the neutral thermal speed and represent the first experimental observations of non-thermal plasma over a range of observing angles, indicating a clear angular dependance. The observed ion temperature anisotropy, deduced by assuming a Maxwellian line-of-sight ion velocity distribution, exceeds the theoretical value for a bi-Maxwellian based on a realistic ion-neutral collision model.

One of the main problems in this study was that the neutral temperature and wind velocity had to be estimated from either the radar data or models, which may introduce a large source of error into the calculations described above. This problem is discussed and the possible consequences are illustrated by comparing the meridional and zonal components of the neutral wind derived from EISCAT observations with those measured simultaneously in the same volume, using a Fabry-Perot interferometer (FPI). A new experiment is planned which will combine low elevation EISCAT observations of the auroral zone/polar cap ionosphere with simultaneous FPI measurements of the neutral wind and temperature, which will address some previously unsolved problems on the collisional and plasma processes that occur in the high-latitude ionosphere/thermosphere system.

## DISCUSSION

**R.Showen**
Were your observations for a monostatic or tristatic configuration?

**Author's Reply**
We have both monostatic and tristatic observations.

**P.Vila**
Do the non-thermal features extend throughout the $F$-region?

**Author's Reply**
Yes. Certainly above about 200 km because collisions will destroy the features at low altitudes.

**R.Schunk**
I would like to caution you about the interpretation of your measurements. If you have 2 km/s ion drifts, you will get ion composition changes and $NO^+$ could be the dominant ion up to 340 km. Therefore, when interpretating the spectra you not only have to worry about bi-Maxwellian features, but you also need to take into account the possibility of both $NO^+$ and $O^+$ in the $F$-region.

**Author's Reply**
I completely agree with your point and that is why we introduced a three species analysis at the end. However, we have a very unusual set of circumstances on this particular occasion and that is why we have $N_2^+$.

**K.Suchy**
What are the numerical values of $\beta_\parallel$ and $\beta_\perp$?

**Author's Reply**
We estimated the values from the data and they are consistent with the accepted models of collisional processes, falling between those obtained for resonant charge exchange and polarization collisions.

# QUIET TIME CONDUCTIVITIES OF THE AURORAL IONOSPHERE

Asgeir Brekke and Chris Hall
The Auroral Observatory
P.O. Box 953
N 9001 Tromsø
Norway

SUMMARY

EISCAT incoherent scatter data of the E- and F region raw electron density profiles for 7 quiet days have been used to derive the auroral ionosphere conductivities and conductances.

Empirical formulas for the solar zenith angle dependence have been obtained which are compared with earlier models and results.

In particular, the Hall- to Pedersen-conductance ratio has been found to be lower during quiet periods than usually reported from similar work on incoherent scatter data and model studies. For disturbed days, however, when the data are corrected for background ionization, this conductance ratio is much higher than models predict.

## 1. THE IMPORTANCE OF THE IONOSPHERIC CONDUCTIVITIES

In the studies of the dynamical coupling between the global ionosphere and the magnetosphere and/or upper atmosphere it is mandatory to have a good knowledge of the ionospheric conductivities. The ionospheric current, $\underline{j}$, for instance, which is a major actor in this coupling process is given by the general Ohm's law as:

$$\underline{j} = \underline{\underline{\sigma}} \cdot \underline{E}' \qquad (1)$$

where $\underline{\underline{\sigma}}$ is the conductivity tensor and $\underline{E}'$ is the electric field with the neutral dynamo field included. All parameters are height dependent.

In a more convenient reference system for ionospheric studies where one axis is along the Earth's magnetic field, $\underline{B}$, this current can be expressed as:

$$\underline{j} = \sigma_P \underline{E}_\perp' + \sigma_H \frac{\underline{B} \times \underline{E}_\perp'}{B^2} + \sigma_\parallel \underline{E}_\parallel' \qquad (2)$$

where $\underline{E}_\perp'$ and $\underline{E}_\parallel'$ are the components of the electric field perpendicular and parallel to $\underline{B}$. $\sigma_P$, $\sigma_H$ and $\sigma_\parallel$ are the height dependent Pedersen, Hall and parallel conductivities, respectively. If, for example, measurements of the electric fields have been obtained, the currents can only be determined when proper values of the conductivities are available. The same also applies of course if $\underline{E}'$ is measured and $\underline{j}$ is to be derived.

From Biot Savart's law we know the relationship between the electric current, $\underline{j}$, and the magnetic perturbation, $\Delta\underline{B}$, this current creates at a distance $\underline{r}$. The relationship is given by:

$$\Delta\underline{B} = \frac{\mu_0}{4\pi} \frac{\underline{j}\underline{k} \times \underline{r}}{r^2} \qquad (3)$$

where $\underline{k}$ is a line element along the current direction and $\mu_0 = 1.26 \times 10^{-6}$ H/m is the permeability constant. For the auroral zone ionosphere where the magnetic field is approximately vertical and the electric field parallel to the magnetic field can be neglected, the ionospheric currents will be mainly horizontal. Furthermore, the current due to the distribution of the electron density will be limited in height and essentially form a sheet current. In this situation it is found that the magnetic deflection on the ground due to an overhead sheet current can be expressed:

$$\Delta\underline{B} = \kappa \left( \int_{z_1}^{z_2} \sigma_P \underline{E}_\perp' \, dz + \int_{z_1}^{z_2} \sigma_H \frac{\underline{B} \times \underline{E}_\perp'}{B^2} \, dz \right) \qquad (4)$$

where $\kappa$ is a proportionality constant depending on the induction in the ground and the geometry of the sheet with respect to the observer. $z_1$ and $z_2$ are the lower and upper borders of the current sheet. If an average electric field within this height region is inferred, Equation (4) can be further simplified:

$$\Delta\underline{B} = \kappa \left( \Sigma_P \underline{E}_\perp + \Sigma_H \frac{\underline{B} \times \underline{E}_\perp}{B^2} \right) \qquad (5)$$

where $\underline{E}_\perp$ now is the average E field perpendicular to $\underline{B}$ in the height region $z_2$ $z_1$, and $\Sigma_P$ and $\Sigma_H$ are the height integrated conductivities (conductances) given by:

$$\Sigma_P = \int_{z_1}^{z_2} \sigma_P \, dz \qquad (6a)$$

$$\Sigma_H = \int_{z_1}^{z_2} \sigma_H \, dz \tag{6b}$$

Observations of magnetic field fluctuations in the auroral zone have been used, since the pioneer work of Birkeland (1), to study equivalent ionospheric current systems. In recent years this type of work has been expanded to also extract informations of the ionospheric electric fields from groundbased magnetic observations (2-4). This is, however, according to Equation (5) not possible unless simultaneous measurements or appropriate models of the Hall- and Pedersen-conductances are available. As currents parallel to the Earth's magnetic field, ideally, do not create any observable effect on ground (5) the neglect of $\underline{E}_{\parallel}$ is not essential to the approach presented above.

In the interaction processes between the magnetosphere and the ionosphere or the lower atmosphere and the ionosphere, the electrical potential distribution plays a fundamental role. It is especially important to understand how electrical potentials varying in space and time propagate downward from the magnetosphere or upward from the mesosphere and lower atmosphere to the ionosphere to interact with the currents there. Schematically we can illustrate this interaction with the flow chart in Figure 1. $U_M$ is the potential created by the magnetospheric dynamo and propagating down to the ionosphere, $U_I$ is the ionospheric potential created by the ionospheric dynamo which can propagate upward towards the magnetosphere, and $U_A$ is the atmospheric potential which may be created in the weather system and propagate up to the ionosphere and even further.

The problem of potential mapping between the three regions can be studied by simultaneously solving the equations:

$$\underline{E} = -\nabla V \tag{7}$$

$$\nabla \cdot \underline{I} = 0 \tag{8}$$

where the ionospheric current is given by Equation (2). As it turns out, the damping of the potentials in space and time depends upon the conductivities (conductances) (6-9). Strongly related to this are also the propagation characteristics of hydromagnetic waves in the ionosphere (magnetic pulsations) which are largely determined by the ionospheric conductivities.

A phenomenon of increasing interest in the polar ionosphere is the forming of auroral arcs and the electrodynamical processes taking place in and around such arcs. In the following we will adopt a very simple model according to Boström (5) of an auroral arc (Figure 2) in order to demonstrate the important role the ionospheric conductances play in this problem.

We assume an infinitely long arc aligned along the Y-axis (Y - positive westward). The width in the north-south direction is a (x - positive north) and the height region covered by the arc is between $z_1$ and $z_2$. The Earth's magnetic field is antiparallel to the vertical axis z (northern hemisphere). Along the arc both outside and within there is an *electric field* $\underline{E}^O$ *directed* westward. The arc is represented in the ionosphere with a region of high conductance, $\Sigma^A$, compared to the conductance, $\Sigma^C$, outside the arc.

In Figure 2 the arc is projected into the x,y plane perpendicular to $\underline{B}$, and $\underline{B}$ is directed into this plane. Inside the arc the electric field $\underline{E}^O$ will drive a Pedersen current given by:

$$I^O_{y,A} = \Sigma^A_P E^O \tag{9a}$$

and a Hall current given by:

$$I^O_{x,A} = \Sigma^A_H E^O \tag{9b}$$

Outside the arc the corresponding currents will be

$$I^O_{y,C} = \Sigma^C_P E^O \tag{10a}$$

$$I^O_{x,C} = \Sigma^C_H E^O \tag{10b}$$

The Hall current inside the arc, however, will accumulate positive charges on one side and negative charges on the other side of the arc. These charges will create an electrical polarization field $\underline{E}^P$ directed along the negative x-axis, which will prevent the charges from accumulating further. This polarization field will drive a Pedersen current inside the arc given by:

$$I^P_{x,A} = -\Sigma^A_P E^P \tag{11a}$$

and a Hall current given by:

$$I^P_{y,A} = \Sigma^A_H E^P \tag{11b}$$

The total current inside the arc is now given by:

$$I_{x,A} = \Sigma_H^A E^O - \Sigma_P^A E^P \tag{12a}$$

$$I_{y,A} = \Sigma_P^A E^O + \Sigma_H^A E^P \tag{12b}$$

and the currents outside the arc is given by:

$$I_{x,C} = \Sigma_H^C E^O \tag{13a}$$

$$I_{y,C} = \Sigma_P^C E^O \tag{13b}$$

As we must have a continuous current across the interface between the arc and the region outside the arc

$$I_{x,A} = I_{x,C} \tag{14}$$

and

$$E^P = (\Sigma_H^A - \Sigma_H^C)E^O/\Sigma_P^A \tag{15}$$

Finally,

$$I_{y,A} = \left[\Sigma_P^A + \frac{\Sigma_H^A}{\Sigma_P^A}(\Sigma_H^A - \Sigma_H^C)\right] E^O \tag{16}$$

Since $\Sigma_H^A$ is usually larger than $\Sigma_P^A$ in an auroral arc and we have assumed that $\Sigma_H^A >> \Sigma_H^C$ the second term in the bracket will be positive and the current along the arc will be enhanced compared to the Pedersen current along the arc driven by $\underline{E}^O$. If we neglect $\Sigma_H^C$ compared to $\Sigma_H^A$ we get

$$I_{y,A} = \left[\Sigma_P^A + \frac{(\Sigma_H^A)^2}{\Sigma_H^A}\right] E^O = \Sigma_P^A \left(1 + \left(\frac{\Sigma_H^A}{\Sigma_P^A}\right)^2\right) E^O \tag{17}$$

and the current enhancement factor inside the arc will be

$$I_{y,A}/I_{y,A}^O = 1 + \left(\frac{\Sigma_H^A}{\Sigma_P^A}\right)^2 \tag{18}$$

From equation (15) we also notice that neglecting $\Sigma_H^C$ compared to $\Sigma_H^A$ is equivalent to

$$\frac{E^P}{E^O} = \frac{\Sigma_H^A}{\Sigma_P^A} \tag{19}$$

and the polarization factor of the electric field as well as the current enhancement factor can be expressed by the ratio $\Sigma_H/\Sigma_P$ inside the arc. The polarization of an auroral arc can therefore according to this model be studied by ion and neutral measurements of the conductances within and outside the arc.

A final example where knowledge of the conductivities are of crucial importance is found in the studies of the Joule or frictional heating associated with ionospheric currents.

The height dependent Joule dissipation rate, $q_J$, due to an ionospheric current given by Equation (2) can be expressed as:

$$q_J = \sigma_P E_\perp'^2 \tag{20}$$

and for the height integrated quantity

$$Q_J = \lambda_P \underline{E}_\perp^2 \tag{21}$$

where $\underline{E}_\perp$ is the average electric field in the height region considered.

In this context it is therefore felt that one of the most important parameters to quantify in a proper manner is the ionospheric conductivity as it plays such an important role in modern studies of the upper atmosphere and ionosphere, especially at high latitudes but also elsewhere.

We find it particularly important to establish a measure of the quiet time auroral ionospheric conductivities in order to understand the full global effects of high latitude disturbances in the upper atmosphere. Only when these quiet time quantities are established and known, can the net effect of the disturbed auroral ionosphere currents to the heating rates of the upper atmosphere be calculated.

## 2. THE ANALYSIS METHOD USED

### 2.1 The ionospheric conductivities

In 1927, Pedersen (10) first pointed out, in connection with radio measurements of the electron density and the dynamo theory, that the geomagnetic field renders the ionospheric electrical conductivity anisotropic. In 1953, Chapman (11), in his review of the electrical conductivity of the ionosphere, gave a theoretical outline for the derivation of this parameter.

The height dependent ionospheric Hall- and Pedersen conductivities which can now be found in standard textbooks on ionospheric physics, can be written as

$$\sigma_H = \frac{N_e \cdot e}{B}\left[\frac{\omega_e^2}{\omega_e^2+\nu_{en}^2} - p_{i1}\frac{\omega_{i1}^2}{\omega_{i1}^2+\nu_{i1n}^2} - p_{i2}\frac{\omega_{i2}^2}{\omega_{i2}^2+\nu_{i2n}^2}\right] \tag{22a}$$

$$\sigma_P = \frac{N_e \cdot e}{B}\left[\frac{\omega_e \nu_{en}}{\omega_e^2+\nu_{en}^2} + p_{i1}\frac{\omega_{i1}\nu_{i1n}}{\omega_{i1}^2+\nu_{i1n}^2} + p_{i2}\frac{\omega_{i2}\nu_{i2n}}{\omega_{i2}^2+\nu_{i2n}^2}\right] \tag{22b}$$

Here plasma neutrality and two different single charged ion species are assumed with the relative contributions $p_{i1}$ and $p_{i2}$ respectively. The other parameters are as follows:

$N_e$ - height dependent electron density
$e$ - electronic charge ($1.6 \times 10^{-19}$ C)
$\omega_e$ - electron gyro frequency ($eB/m_e$)
$\omega_i$ - ion gyro frequency ($eB/m_i$)
$\nu_{en}$ - electron neutral collision frequency
$\nu_{in}$ - ion neutral collision frequency.

For the magnetic field, B, the IGRF 65 model (12) has been used. The two ion species are assumed to have ion masses $m_{i1} = 30.5$ and $m_{i2} = 16$, respectively. For the lighter ion mass this corresponds to atomic oxygen while the heavier ion mass represents a mixture of 75% $NO^+$ and 25% $O_2^+$.

For the relative abundance of the light ion ($O^+$) and the heavy ion ($NO^+$, $O_2^+$ mixture) the model presented in Figure 4 from Lathuillère (13) has been adopted. This model, however, is static and independent of time of day or season.

The height integrated values that will be presented are obtained between the heights 90 and 330 km respectively.

For the collision frequencies we have adopted the following models:

$$\begin{aligned}\nu_{en} = {} & 2.33 \times 10^{-17} \times n(N_2) \times (1 - 1.21 \times 10^{-4} \times T_n) \times T_n \\ & + 1.82 \times 10^{-16} \times n(O_2) \times (1 + 3.6 \times 10^{-2} \times \sqrt{T_n}) \times \sqrt{T_n} \\ & + 8.9 \times 10^{-16} \times n(O) \times (1 + 5.7 \times 10^{-4} \times T_n) \times \sqrt{T_n}\end{aligned} \tag{23}$$

according to Schunk and Nagy (14) when $T_n = T_e$.

$$\nu_{i1n} = \left[4.23 \times n(N_2) + 4.18 \times n(O_2) + 2.38 \times n(O)\right] \times 10^{-16} \tag{24a}$$

$$\begin{aligned}\nu_{i2n} = {} & \left[6.82 \times n(N_2) + 6.66 \times n(O_2)\right. \\ & + 3.42 \times n(O) \times \sqrt{T_n} \times (1.08 - 0.139 \times \log_{10} T_n \\ & \left. + 4.51 \times 10^{-3} \times (\log_{10} T_n)^2)\right] \times 10^{-16}\end{aligned} \tag{24b}$$

from Schunk and Walker (15) when $T_r = (T_n+T_e)/2 = T_n$. The neutral densities are in units of $m^{-3}$ taken from the MSIS model (16).

In this study we apply the raw electron density profiles obtained by the EISCAT CP1 (17) program for 7 quiet days in 1985 to 1987. The CP1-program yields ionospheric power profiles sampled at every 1.5 km with a pulse length of 60 μs. The profiles are derived between altitudes 90 and 330 km at an elevation angle 76.5° to the south, i.e. approximately up along the field line from the Tromsø EISCAT site. The on-line integration time is 10 sec, however, the post integration time used in the data presentation is 5 mins.

The raw densities are related to the true electron densities by a correction factor given by (17,18):

$$\frac{N_e}{N_{raw}} \propto \left(1 + \frac{T_e}{T_i} + \alpha^2\right)(1 + \alpha^2) \tag{25}$$

accounting for unequal ion ($T_i$)- and electron ($T_e$) temperature and Debye length effects. Since the part of the electron density profile that will contribute most to the conductivities will be situated below 180 km it is not believed that $T_e$ and $T_i$ will differ substantially from each other in these quiet periods that are our main concern. Above 180 km, however, the difference between $T_e$ and $T_i$ increases gradually with height, and therefore the raw electron density used in this work will be an underestimate. As will be shown below, the contribution from the electron densities at these altitudes to the conductances will be minor, and this underestimation therefore represents a negligible error source, 20% at most.

The Debye length correction for the EISCAT radar (UHF) used in this work is given approximately by (19):

$$\alpha \approx 7.3 \times 10^{6} \frac{T_e}{N_e} \qquad (26)$$

For E-region electron densities greater than a few times $10^{10}$ $m^{-3}$ and usual quiet time E-region temperature of 300 to 600 K, this correction is less than 10%. Figure 5 is a flow chart showing the principles for deriving the conductances.

2.2 Presentation of electron density- and conductivity profiles

In order to illustrate the dynamics of the auroral ionosphere conductivities, we present in Figure 6 a set of electron density profiles together with Hall- and Pedersen conductivity profiles derived according to Equation (22) for four different time periods on Aug. 6-7, 1985.

In none of the examples shown is it likely that the contribution to the Pedersen conductance from the height region above 180 km can be significant.

3. PRESENTATION OF CONDUCTANCES

In Figure 7 we show the Hall- and Pedersen conductances derived as functions of time for the whole 24 hour EISCAT experiment on Aug. 06-07, 1985 (LT = UT + 1 hr). The curves are remarkably regular in their shape except for a few events clearly due to auroral particle precipitation. They reveal a smooth daily variation of the conductances related to the change in solar zenith angle. The scatter in the measurements during the daytime hours 0700-1800 UT is so small that, except for a few erratic points, it is within the thickness of the line.

In the bottom panel of Figure 7 we also present the ratio $\Sigma_H/\Sigma_p$ which for the quiet periods on this day is close to 1.3. For the disturbed periods, however, this ratio reaches values close to 5. Since these events occur at a time when the background conductances caused by solar radiation are substantial, the ratio given in Figure 7 is not a measure of the auroral particle precipitation alone.

In Figure 8 we present the Hall- and Pedersen conductances as functions of the solar zenith angle for these periods in Aug. 6-7, 1985, where auroral particle precipitation is not present. Both conductances clearly decrease with increasing solar zenith angle.

Also included in Figure 8 are the least squares fit to each set of conductances. The empirical formula used in the fit for both sets is:

$$\Sigma = a \cos\chi + b(\cos\chi)^{1/2} + c \qquad (27)$$

The rationale for using this approach will be discussed below, evidently however, Equation (27) represents a very good fit to both data sets. Table I gives the a, b and c factors for each of the curves in Figure 8.

In Figure 9 we present Hall and Pedersen conductances derived for a 24 hour experiment on June 16-17, 1987. This data set is more disturbed than the previous example. This is in particular clear from the strong conductance enhancement occurring between 1900 UT on June 16 and 0200 UT on June 17. Note that the scale to the left in the upper panel is increased by a factor of 2 compared to that in Figure 7. The conductances observed in the relatively quiet period between 1200 and 1800 are also more variable than for the other day, indicating that some particle precipitation may be present even at this time.

The quiet time conductance ratio $\Sigma_H/\Sigma_p$ shown in the middle panel of Figure 9 is about 1.3 for this day. During the disturbed event at midnight, however, this ratio reaches peak values of 2.80. Again is this not a true measure of the effect of the auroral precipitation because the ionization background due to solar radiation is substantial during part of this event. In the lower panel of Figure 9 we show this ratio when corrected for background ionization. These results, however, will be especially dealt with below.

Figure 10 shows the Hall- and Pedersen conductances for those periods where particle precipitation is not strongly present, as function of the solar zenith angle. The date is the same as for the data shown in Figure 9. Again the solar zenith angle relationship is evident. The scatter of the points is somewhat larger than for the previous day shown, probably because some data points do include auroral particle precipitation. The empirical functions obtained by a least squares fit to Equition (27), however, are very similar to those found for the other days as shown in Table I.

4. DISCUSSIONS

4.1 The solar zenith angle dependence

Since the quiet daytime values of the conductances in Figure 7 and 9 are so well behaved over periods as long as 10 hours, we have used the quiet time sequences, when no particle precipitation is present, to derive the empirical solar zenith angle dependence on the conductances. In Figure 8 and 10 these functions are shown for the Hall- and Pedersen conductances for the 2 EISCAT experiments discussed here, and in

Table I are listed the coefficients a, b and c as derived from 7 different data sets.

Adopting the continuity equation for the ionospheric electrons we have

$$\frac{dN_e}{dt} = q_{solar} - L_e - \mathrm{div}(N_e \cdot \underline{v}) \tag{28}$$

Here $q_{solar}$ is the electron production due to the solar EUV radiation and $L_e$ is the electron loss term.

Assuming a steady state, justifiable for the quiet time analysis we are performing here, the loss rate will be equal to the production rate at all times, and

$$L_e = q_{solar} \tag{29}$$

The production maximum at a given solar zenith angle due to solar radiation for an atmosphere with constant scale height, can be expressed as

$$q_{solar} = q_o \cos\chi \tag{30}$$

where $\chi$ is the solar zenith angle and $q_o$ is the production for overhead sun. In the lower E-region we can expect that dissociative recombination is the dominant loss mechanism such that

$$L_e = \alpha N_e^2 \tag{31}$$

and the electron density or E-region ionization maximum is given by:

$$N_e \propto (\cos\chi)^{1/2} \tag{32}$$

In the upper E-region and the lower F-region where $O^+$ is rapidly converted to molecular ions ($O_2^+$ and $NO^+$) the loss can be expressed as

$$L_e = \beta N_e \tag{33}$$

or the electron density at F-region ionization maximum is given by:

$$N_e \propto \cos\chi \tag{34}$$

In the integration process we are actually adding electron densities which partly follow the $\alpha$ and partly the $\beta$ decay mechanisms. Although Equations (32) and (34) apply for the E-region and F-region maxima only it is reasonable to assume that the height integrated quantities will to a first approximation be a function of both these two terms.

We therefore try to make a least squares fit of the data presented in Figure 7 and 9 to the function combining Equations (32) and (34) as given by Equation (27).

It is evident from Table I that for all cases we have studied so far the factor c is very close to zero and can be neglected. In the bottom line of Table I we also derive the mean values and the standard deviations on these mean values for each of the sets of constants a and b, based on the 7 days of data.

4.2 *Comparison with other results*

We will now use the empirical formula for $\Sigma_p$ derived in this work

$$\Sigma_p = (3.05 \pm 0.85)\cos\chi + (4.06 \pm 0.78)(\cos\chi)^{1/2} \tag{35}$$

to compare our results with similar results derived from empirical formulas (see Table II) obtained by other authors [19-25]. The results of this comparison are shown in Figure 11 where $\Sigma_p$ is plotted versus the solar zenith angle for each of the models. Also indicated are the corresponding standard deviations on each point of our measurements.

For the experimental work [19] on EISCAT we notice that we have good correspondence for solar zenith angles lower than 40 degrees. As very few measurements exist for such low angles at EISCAT our formulas are essentially empirical extrapolations at these angles. In the region where we have observations at EISCAT ($\chi > 40^o$) our two results disagree. This is probably due to the fact that Schlegel [19] has made no attempt to establish the quiet time solar zenith angle dependence of the conductances but rather estimated an empirical function through the mean of the data set covering disturbed as well as quiet days. Therefore the conductances derived by Schlegel [19] will be higher than our quiet time values because his data will be contaminated by auroral particle precipitation.

The experimental Chatanika work by Robinson and Vondrak [23] is based on 13 days of observations at varying solar flux between 70 and 221 with an average value of 136 ($\times 10^{-22}$ W/m$^2$ Hz). The data were obtained between 1972 and 1981. Their results are in fairly good agreement with ours. In the empirical formula of Robinson and Vondrak [23] (see Table II) we have used $S_a$ = 81 ($\times 10^{-22}$ W/m$^2$ Hz) which is the mean value in our data set. Robinson and Vondrak [23], however, do point out that their empirical fit to the conductivity versus solar zenith angle relationship is systematically offset to lower values of the solar flux below 140.

The results obtained by Vickrey et al. (21) based on one day of data are too low compared to our observations. The reason for this is not clear. The empirical formula derived by Mehta (20), also made of one day of data only, agree rather well with our formula for $\Sigma_P$ versus $\chi$. This, however, is rather fortuitous as Mehta (20) used model collision frequencies that were a factor of 2 higher than the one used by Vickrey et al. (21) and in the present work. With a more proper choice of collision frequency model, however, Mehta's result would be in better agreement with (21).

de la Beaujardière et al. (22) also used collision frequencies that were too high by a factor of 2 and they obtained values for $\Sigma_P$ that are almost twice as high as the values presented in this work. Another reason for this high value is probably the strong solar flux ($S_a$ = 156) present for the day analysed by de la Beaujardière et al. (22) compared to the average solar flux of 81 for our data.

Rasmussen et al. (24) obtained the following empirical formula for the Pedersen conductance based on 13 days of observations by the Chatanika radar:

$$\Sigma_P = \frac{4.5}{B} \times (1 - 0.85*v^2) * (1 + 0.15*u + 0.05*u^2) \qquad (36)$$

where B is a factor depending on the magnetic field strength, $v = \chi/90^o$ where $\chi$ is the solar zenith angle and $u = S_a/90$ where $S_a$ is the 10.7 cm Ottawa 2800 MHz solar flux index. In Figure 11 we have drawn the function given by Equation (23) for B = 0.53 corresponding to the Tromsø value of the magnetic B field (~ 53000 nT) and $S_a$ = 81 which is equivalent to the mean value of this parameter for the days listed in Table II. For all solar zenith angles our Pedersen conductance values are lower than those derived by Rasmussen et al. (24). Although their formula is derived for $S_a$ between 70 and 250 with a mean value of 144, only 3 days in their data set of 13 days exhibit $S_a$ values below 100. For our data set of 7 days no day has an $S_a$ value above 100 and the average value is 81. It is then possible that the formula obtained by Rasmussen et al. (24) is biased toward high $S_a$ values and overestimates the Pedersen conductances for low $S_a$-values.

In order to study the relative importance of the two terms $a\cos\chi$ and $b(\cos\chi)^{1/2}$ in Equation (27) for the data sets presented in Table I, the relative magnitude $a/(a+b)$ in per cent is also shown for each of the conductances in Table I.

As pointed out above, the electron density below 140 km will contribute most strongly to the Hall conductance. In this height regime the dissociative recombination process is believed to be most important, and the quadratic loss term (Equation (31)) is expected to be most dominant. It is therefore surprising to find that for 6 of the 7 cases studied so far it is the linear loss term that is most dominant to the Hall conductance according to our empirical fits to Equation (27). Only for one day (March 24, 1987) is the contribution from the linear loss term less than 50% (35%), while it is larger than 70% for 4 days.

Also shown in Table I are the similar ratios $(a_P/(a_P+b_P))$ for the Pedersen conductance, and it is found that in the present data set studied, this ratio is on the average 42%, for 3 days it is above 50%, and for 2 days it is below 40%. For the Pedersen conductance therefore it appears that the quadratic and linear loss terms contribute with approximately the same amount. As the Pedersen conductance gets about equal contribution from the altitude region above and below 120 km according to the results presented here, it is expected that the two terms will contribute more equally to Equation (27) for the Pedersen- than for the Hall conductance.

4.3 The conductivity ratio

Since the work by Brekke et al. (26) the conductance ratio has been used extensively as an indicator of the energetic particle precipitation (19 21, 24, 27 and 28). Several authors (20, 24 and 28) have also derived models for how this ratio will change with the characteristic energy of the precipitating particles. When the conductance ratio derived from observed electron density profiles is used as an indicator of the energy of the precipitating particles, it is important, however, to first correct for the contributions to the individual conductances from solar radiation. It is well known that when a precipitation event occurs in the morning hours, the energy of the particles is usually higher than for a precipitating event in the evening sector. If, however, the conductances are large in the morning hours because of solar EUV radiation, a precipitation event will only contribute relatively weakly to an already enhanced ratio. To get a better indicator of the influence of particle precipitation, the background conductivity must therefore be subtracted.

The ratio relevant for the precipitating events is then

$$\frac{\Delta\Sigma_H}{\Delta\Sigma_P} = \frac{\Sigma_H - \Sigma_H^o}{\Sigma_P - \Sigma_P^o} \qquad (37)$$

where the $\Sigma_H^o$ and $\Sigma_P^o$ are the background conductances which must be subtracted.

With this approach, however, we introduce a new problem during quiet times as $\Delta\Sigma_H$ and $\Delta\Sigma_P$ will be close to zero. If the background model conductances we subtract are good estimates, the net result in Equation (37) will be poorly defined.

This problem is illustrated in the bottom panel of Figure 9 where the conductance ratio $(\Delta\Sigma_H/\Delta\Sigma_P)$ for June 16-17, 1987 is calculated by subtracting the empirical functions derived from the original data set for this day according to Table I. Before 1900 UT there are several points that are scattered. Some of these points may be real because of possible presence of auroral particle precipitation before 1800 UT as already mentioned, but most of them are simply due to the problem of dividing very small terms with each other after forming $\Delta\Sigma_H$ and $\Delta\Sigma_P$. When $\Delta\Sigma_H$ or $\Delta\Sigma_P$ is negative the ratio $\Sigma_H/\Sigma_P$ is used in the bottom panel of Figure 9. For the precipitation event taking place between 1900 UT on June 16, 1987 and 0300 UT on

June 17, 1987 the importance of subtracting the background is clear. When comparing the ratio $\Delta\Sigma_H/\Delta\Sigma_P$ which reaches a maximum value of ~ 7.5 at 0300 UT in the lower panel of Figure 9 with the ratio $\Sigma_H/\Sigma_P$ reaching a maximum value ~ 3 in the middle panel of Figure 9, it is clear that the conductance ratio without correction for background ground ionization is meaningless as far as the energy of the precipitating particles is concerned. At the first maximum observed at 2025 UT $\Sigma_H/\Sigma_P$ is equal to ~ 2.2 while $\Delta\Sigma_H/\Delta\Sigma_P$ is equal to ~ 2.5, thus indicating that the background ionization at this time of day is almost negligible in this respect.

The morning event between 0000 UT and 0300 UT is certainly characterized by particles of increasing energy as the ratio $\Delta\Sigma_H/\Delta\Sigma_P$ increased from 1.5 to 7.5. Contrary to the work by Vickrey et al. (19) and Rasmussen et al. (24) who only obtain maximum values of $\Sigma_H/\Sigma_P$ of about 4 when the conductances are produced by auroral particles only, we find ratios as high as 7.

## 5. CONCLUSION

We have, based on 7 relatively quiet summer days, presented a study of ionosphere Hall- and Pedersen conductivities and conductances, and found that the quiet time conductances (height integrated quantities) are well behaved functions of the solar zenith angle as long as no auroral particle precipitation is present.

We have used the so-called raw electron density profiles obtained by the EISCAT-CP1 experiment. The raw electron densities are proportional to the returned power, but no corrections are made for differences in the ion- and electron temperatures, the Debye correction has also been omitted in this work. As the ion- and electron temperature differs only substantially on quiet days above say 180-200 km it is not expected that neglecting this temperature difference will have a significant influence on our results. We have demonstrated on the contrary that the electron densities above 180 km only contributes 30% at most to the total Pedersen conductance and nothing at all to the Hall conductance. Neither do we believe that neglecting the Debye length correction is of significant importance since this is normally effective below 90 km and our integration window is taken from 90 km and up.

Our results have been compared with similar empirical formulas presented by other authors. Discrepancies found between the different formulas are most probably related to differing choice of collision frequency model. Other reasons for some of the discrepancies are varying solar flux and improper choice of quiet day. In the auroral zone, where auroral particle precipitation often is present, extra care must be taken in order to discriminate periods with precipitation.

We stress the need to subtract the background ionization due to solar radiation, when using values for the conductance ratio to infer the hardness of the precipitating auroral particles. It is shown that after a proper correction for this background ionization, the characteristic ratio ($\Delta\Sigma_H/\Delta\Sigma_P$) for auroral precipitation events is enhanced by a factor of two up to values of 7 or more. These are much higher than those presented in earlier theories.

The empirical formulas obtained for the conductances do include two terms which are related to a linear electron loss rate ($\cos\chi$) and a quadratic electron loss rate ($\sqrt{\cos\chi}$). Surprisingly enough, the linear loss rate is found on the average to contribute relatively more to the Hall conductance (78%) than to the Pedersen conductance (42%). We are at the moment not able to give any explanation for this finding. As the electron density in the lower E-region, where the dissociative recombination is the main electron loss mechanism, contributes most to the Hall conductance we would expect the quadratic electron loss term to be the dominant one.

ACKNOWLEDGEMENT

We are indepted to Dr. Peter Stauning for enlightening discussions and to Mrs. Liv Larssen for preparing manuscript and figures. We also wish to thank the EISCAT staff. EISCAT is being jointly funded by the Science and Research Council (U.K.), Centre National de la Recherche Scientifique (France), Max-Planck Gesellschaft (F.R.G.), Suomen Akatemia (Finland), Norges Almenvitenskapelige Forskningsråd (Norway), and Naturvetenskapliga Forskningrådet (Sweden). This project has been supported by grants from Norges Almenvitenskapelige Forskningsråd.

REFERENCES

1. BIRKELAND, K., 1901, "Expédition Norvégienne pour l'étude des Aurores Boréales. Resultats des Recherce Magnétiques", Christiania.

2. KAMIDE, Y., A.D. RICHMOND, and S. MATSUSHITA, 1981, "Estimation of Ionospheric Electric Fields, Ionospheric Currents and Field-Aligned Currents From Ground Magnetic Records", J. Geophys. Res., 86, 801-813.

3. AHN, B.-H., R.M. ROBINSON, Y. KAMIDE, and S.-I. AKASOFU, 1983, "Electric conductivities, electric fields and auroral particle energy injection rate in the auroral ionosphere and their empirical relations to the horizontal magnetic disturbances", Planet. Space Sci., 31, 641-653.

4. AHN, B.-H., S.-I. AKASOFU, and Y. KAMIDE, 1983, "The Joule heat production rate and the particle energy injection rate as a function of the geomagnetic indices AE and AL", J. Geophys. Res., 88, 6275-6287.

5. BOSTRØM, R., 1977, "Current systems in the magnetosphere and ionosphere", in Radar Probing of the Auroral Plasma, ed. A. Brekke, Universitetsforlaget, Tromsø, Oslo, Bergen, 257-284.

6. REID, G.C., 1965, "Ionospheric Effects of Electrostatic Fields Generated in the Outer Magnetosphere", Radio Sci., 69, 827-837.

7. BOSTRØM, R., 1973, "Electrodynamics of the Ionosphere", in Cosmical Geophysics, ed. A. Egeland, Ø. Holter and A. Omholt, Universitetsforlaget, Oslo.

8. PARK, C.G., 1976, "Downward Mapping of High-Latitude Ionospheric Electric Fields to the Ground", J. Geophys. Res., 81, 168-174.

9. PARK, C.G. and M. DEJNAKARINTRA, 1973, "Penetration of Thundercloud Electric Fields into the Ionosphere and Magnetosphere, 1. Middle and Subauroral Latitudes", J. Geophys. Res., 78, 6623-6633.

10. PEDERSEN, P.O., 1927, "The Propagation of Radio Waves", Danm. Naturvid. Samf., A, nr. 15a, Copenhagen.

11. CHAPMAN, S., 1953, "The Electrical Conductivity of the Ionosphere: A Review", Nuovo Cimento, 4, 1385-1412.

12. AKASOFU, S.-I. and S. CHAPMAN, 1972, "Solar Terrestrial Physics", Oxford, p. 80.

13. LATHUILLÈRE, C., 1985, Private communication.

14. SCHUNK, R.W. and A.F. NAGY, 1978, "Electron Temperature in the F Region of the Ionosphere: Theory and Observations", Rev. Geophys. Space Phys., 16, 355-399.

15. SCHUNK, R.W. and J.C.G. WALKER, 1973, "Theoretical Ion Densities in the Lower Ionosphere", Planet. Space Sci., 21, 1875-1896.

16. HEDIN, A.E., 1983, "A Revised Thermospheric Model Based on Mass Spectrometer and Incoherent Scatter Data: MSIS-83", J. Geophys. Res., 88, 10170-10188.

17. RISHBETH, H. and P.J.S. WILLIAMS, 1985, "The EISCAT Ionospheric Radar: the System and its Early Results", Q. Fr. R. astr. Soc., 26, 478-512.

18. FARLEY, D.T., 1966, "A theory of incoherent scattering of radio waves by a plasma", J. Geophys. Res., 71, 4091-4098.

19. SCHLEGEL, K., 1987, "Auroral zone E-Region conductivities during minimum derived from EISCAT data", Submitted to J. Atmos. Terr. Phys.

20. MEHTA, N.C., 1978, "Ionospheric electrodynamics and its coupling to the magnetosphere", Ph. D. thesis, Univ. of California San Diego.

21. VICKREY, J.F., R.R. VONDRAK, and S. MATHEWS, 1981, "The Diurnal and Latitudinal Variation of Auroral Zone Ionospheric Conductivities", J. Geophys. Res., 86, 65-75.

22. de la BEAUJARDIÈRE, O., M.J. BARON, V.B. WICKWAR, C. SENIOR, and J.V. EVANS, 1982, "MITHRAS: A Program of Simultaneous Radar Observations of the High-Latitude Auroral Zone", Contract report F49620-81-C-0042, SRI International, Menlo Park, California.

23. ROBINSON, R.M. and R.R. VONDRAK, 1984, "Measurements of E Region Ionization and Conductivity Produced by Solar Illumination at High Latitudes", J. Geophys. Res., 89, 3951-3956.

24. RASMUSSEN, C.E., R.W. SCHUNK, and V.B. WICKWAR, 1987, "A Photochemical Equilibrium Model of Ionospheric Conductivity", Submitted to J. Geophys. Res.

25. SENIOR, C., 1980, "*Les conductivités ionosphérique et leur rôle dans la convection magnétosphérique: Une étude expérimentale et théorique*", Diplome de docteur de 3e cycle, L'Univ. Pierre et Marie Curie, Paris.

26. BREKKE, A., J.R. DOUPNIK, and P.M. BANKS, 1974, "Incoherent Scatter Measurements of E-Region Conductivities and Currents in the Auroral Zone", J. Geophys. Res. 79, 3773-3790.

27. FULLER-ROWELL, T.J. and D.S. EVANS, 1987, "Height Integrated Pedersen and Hall Conductivity Patterns Inferred From the TIROS-NOAA Satellite Data", J. Geophys. Res., 92, 7606-7618.

28. ROBINSON, R.M., R.R. VONDRAK, K. MILLER, T. DABBS, and D. HARDY, 1987, "On Calculating Ionospheric Conductances From the Flux and Energy of Precipitating Electrons", J. Geophys. Res., 92, 2565-2569.

TABLE I. The coefficients a, b and c obtained by fitting the different data sets to Equation (27). Also indicated are the relative importance in per cent of the linear loss term Equation (33) to the different conductances. The solar flux $S_a$ index for the different days is also presented.

| Date | $\Sigma_P$ | | | | $\Sigma_H$ | | | | $S_a^{(3)}$ |
|---|---|---|---|---|---|---|---|---|---|
| | $a_P$ | $b_P$ | $c_P$ | $a_P^{(1)}$(%) | $a_H$ | $b_H$ | $c_H$ | $a_H^{(2)}$(%) | |
| Aug. 6, 1985 | 4.20 | 3.08 | 0.00 | 58 | 8.24 | 1.66 | 0.00 | 83 | 78 |
| Sept. 3,4, 1985 | 1.75 | 4.62 | -0.01 | 28 | 4.10 | 3.99 | -0.01 | 51 | 73 |
| July 15, 1986 | 2.80 | 3.95 | -0.01 | 42 | 7.15 | 2.46 | -0.03 | 74 | 74 |
| March 24, 1987 | 2.50 | 4.40 | -0.01 | 36 | 2.54 | 4.74 | -0.01 | 35 | 76 |
| April 14, 1987 | 3.89 | 3.61 | 0.00 | 52 | 6.51 | 2.32 | 0.00 | 74 | 99 |
| May 12, 1987 | 4.39 | 3.44 | 0.00 | 56 | 8.86 | 1.54 | -0.01 | 85 | 85 |
| June 16, 1987 | 1.81 | 5.35 | 0.00 | 25 | 6.59 | 3.26 | 0.00 | 67 | 81 |
| Average values | 3.05 ± 0.85 | 4.06 ± 0.78 | | 42 ± 13 | 6.28 ± 2.24 | 2.86 ± 1.20 | | 78 ± 14 | 81 ± 9 |

(1) actually $a_P/(a_P+b_P)$ in %.

(2) actually $a_H/(a_H+b_H)$ in %.

(3) The daily 10.7-cm solar radio flux at Ottawa in units of $10^{-22}$ W/m$^2$ Hz (adjusted to 1 AU).

TABLE II. Empirical formulas derived for $\Sigma_H$ and $\Sigma_P$ by different authors on the basis of incoherent scatter data since 1972. Also indicated in the last column are the average solar flux values for the days of observation.

| | Empirical formulas | | Data base | |
|---|---|---|---|---|
| | $\Sigma_P$ | $\Sigma_H$ | Dates | $\overline{S}_a$ [1] |
| Mehta; 1978 (20) | $7.1 * (\cos\chi)^{0.44}$ | $13.7 * (\cos\chi)^{0.45}$ | May 24 25, 1972 | 136 |
| Senior; 1980 (25) | $9.6 * \cos\chi + 1.6$ | $15.8 * \cos\chi + 2.3$ | June 12, 1978 | 116 |
| Vickrey et al.; (21) | $5 * (\cos\chi)^{1/2}$ | $10 * (\cos\chi)^{1/2}$ | April 6, 1977 | 76 |
| de la Beaujardière et al.; 1982 (22) | $10 * \cos\chi + 2$ | $16 * \cos\chi + 3$ | June 10-11, 1981 | 156 |
| Robinson and Vondrak; 1984 (23) | $0.88 * \sqrt{S_a * \cos\chi}$ | $1.5 * \sqrt{S_a * \cos\chi}$ | 13 days from 1972 to 1981 | 136 |
| Schlegel; 1987 (19) | $6.4 * (\cos(\chi-12^\circ))^{0.54}$ | | 27 days in 1985-86 | |
| Rasmussen et al.: 1987[2] (24) | $\frac{4.5}{B} * (1 - 0.85v^2) * (1 + 0.15u + 0.05u^2)$ | $\frac{5.6}{B} * (1 - 0.9v^2) * (1 + 0.15u + 0.05u^2)$ | 13 days from 1976 to 1981 | 144 |

1 The daily 10.7-cm solar radio flux at Ottawa in units of $10^{-22}$ W/m$^2$ Hz (adjusted to 1 AU).

2 B is the magnetic field strength given in units of Wb/m$^2$.

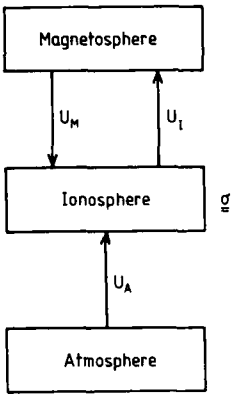



Figure 1 Flow diagram showing the principles of mapping electric potentials between the magnetosphere, ionosphere and atmosphere, and also illustrating the importance of the ionospheric conductivities in this coupling process.

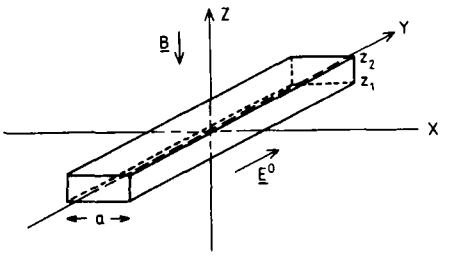



Figure 2 Simple model of an auroral arc represented by an infinite strip of enhanced conductivities between the heights $z_1$ and $z_2$. The width of the strip is a. The Earth's magnetic field, $\underline{B}$, is perpendicular to the arc while an homogeneous electric field, $\underline{E}_0$, is pointing along the arc, both outside and within the arc.

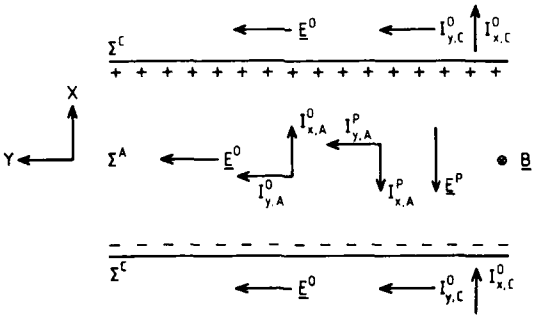



Figure 3 The principles for polarization in and near an auroral arc. See the text for explanation.

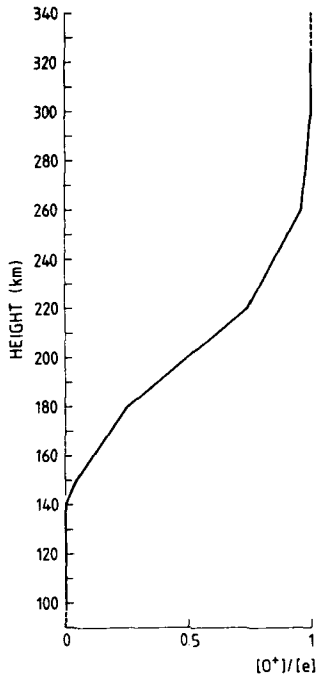



Figure 4 The relative abundance of atomic oxygen ions as function of altitude used in this paper. (Courtesy to Lathuillère (13).)

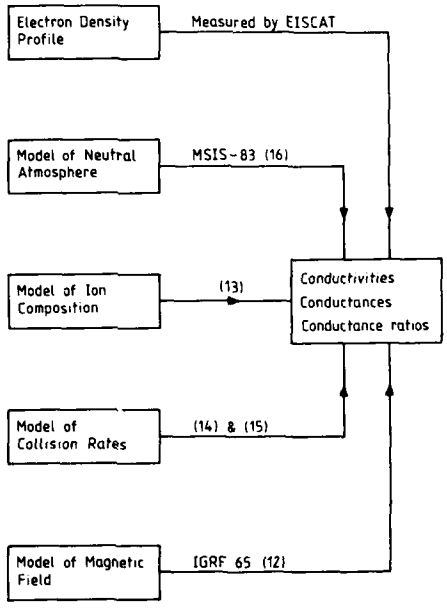



Figure 5 A flow diagram showing the principles for deriving the ionospheric conductivities and conductances.

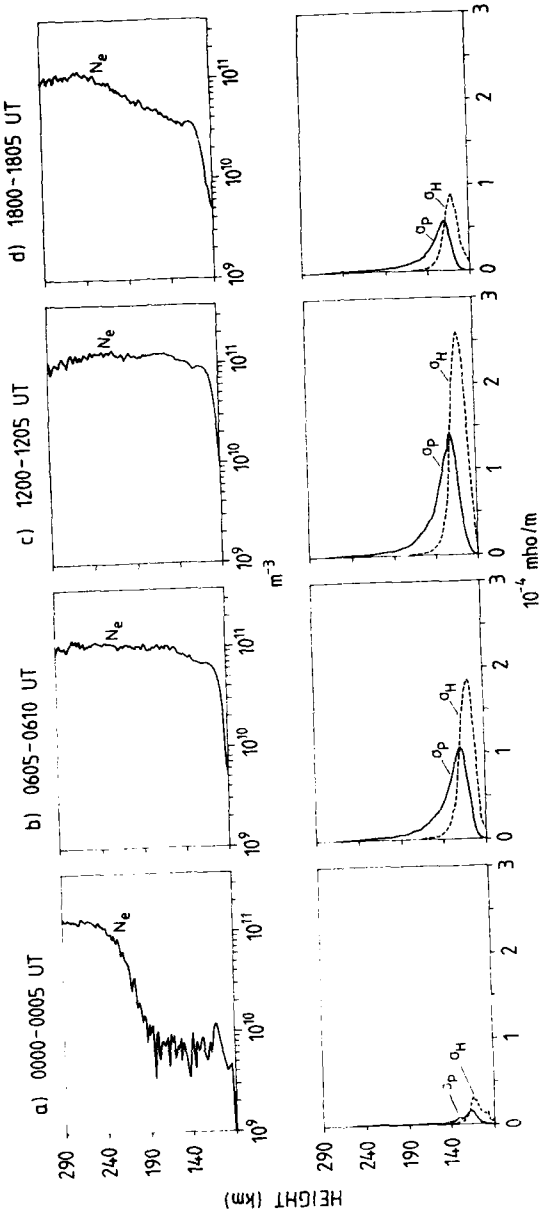



Figure 6 Upper panel: Four characteristic electron density profiles for midnight, morning, midday and evening conditions on Aug. 6/7, 1987.
Lower panel: Corresponding Hall- and Pedersen conductivity profiles.

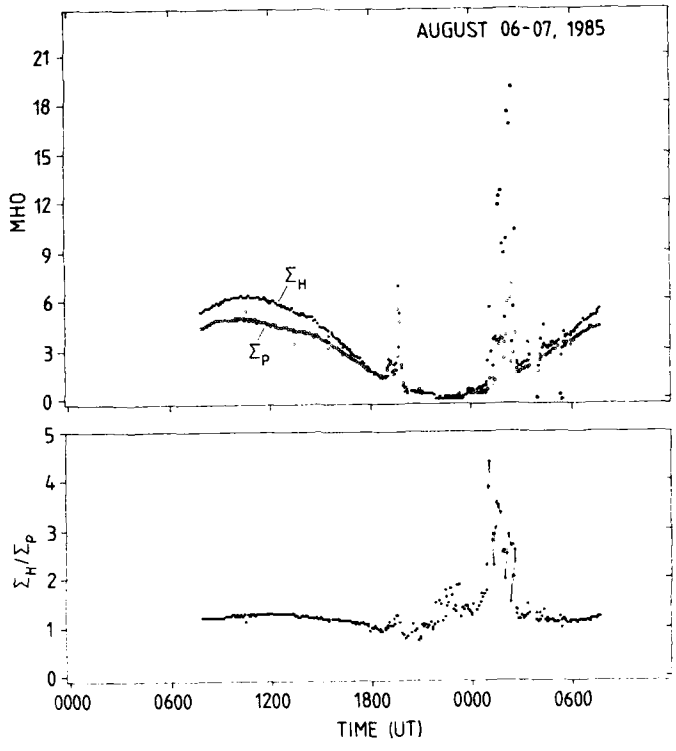



Figure 7 *Upper panel: Hall- and Pedersen conductances as function of time for an EISCAT* experiment conducted in Aug. 06-07, 1985.
Lower panel: The corresponding conductance ratio ($\Sigma_H/\Sigma_P$) as function of time.

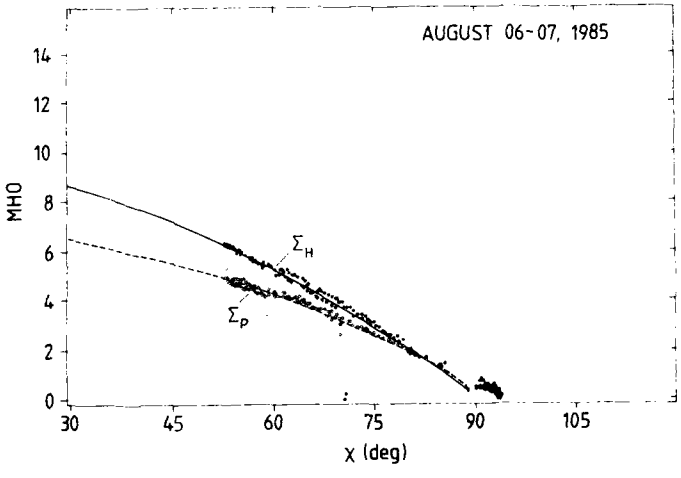



Figure 8 Hall- and Pedersen conductances when periods of auroral particle precipitation is excluded as function of solar zenith angle in degree for the experiment conducted on Aug. 06-07, 1985.

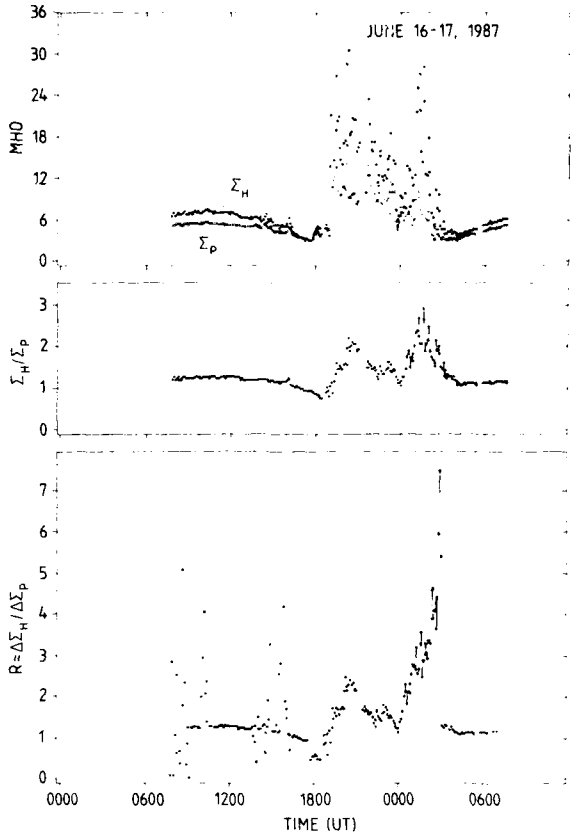



Figure 9 Upper panel: Same as Figure 7 except for the experiment conducted on June 16-17, 1987. Note that resolution on the vertical scale is twice the resolution on Figure 7.
Middle panel: Hall- and Pedersen conductance ratios ($\Sigma_H/\Sigma_P$) as function of time for the experiment conducted on June 16-17, 1987.
Lower panel: Corrected Hall- and Pedersen conductance ratios ($\Delta\Sigma_H/\Delta\Sigma_P$) for the experiment conducted on June 16-17, 1987. The background ionization due to solar radiation is subtracted. When either $\Delta\Sigma_H$ or $\Delta\Sigma_P$ is negative the $\Sigma_H/\Sigma_P$ values are used.

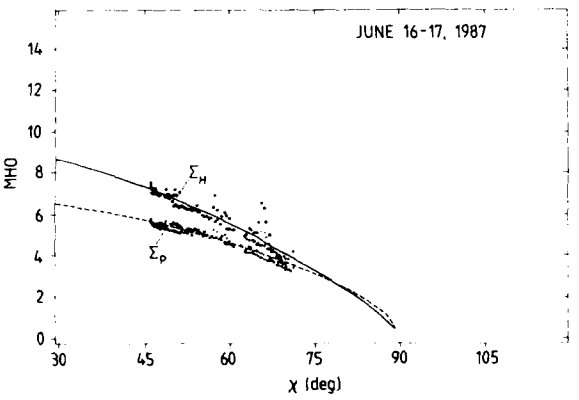



Figure 10 Same as Figure 8 except for the experiment conducted on June 16-17, 1987.

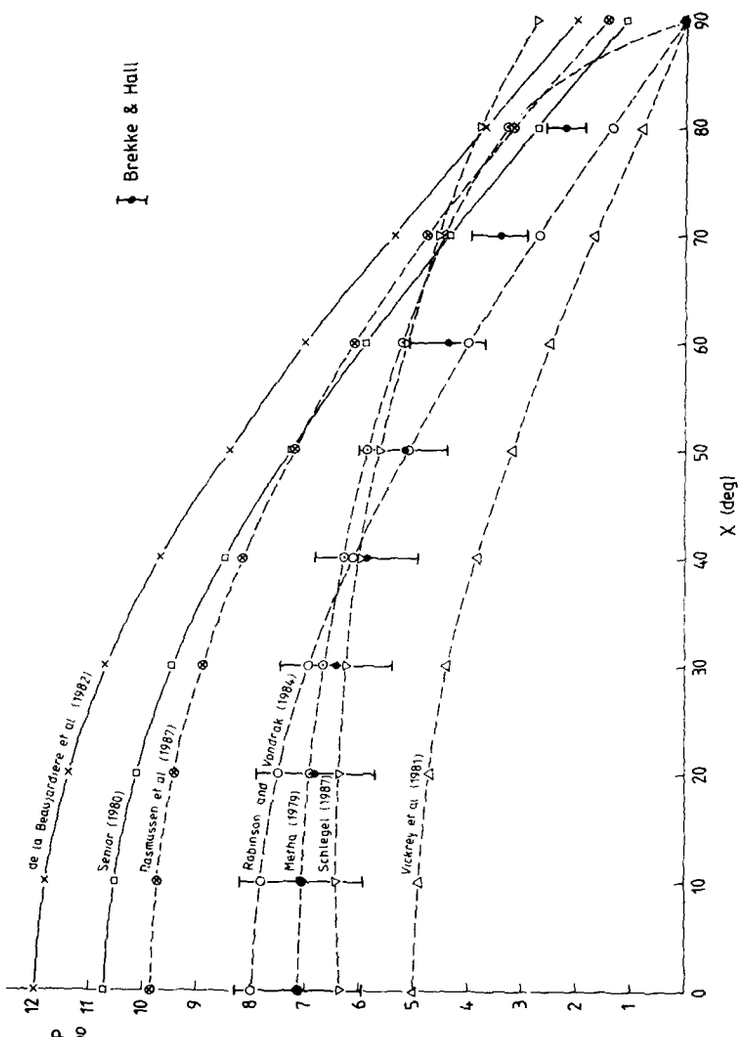



Figure 11 Comparison between the empirical formula presented for the Pedersen conductance versus solar zenith angle in degrees in this work with similar empirical formulae derived by other authors. (See Table II and the text for discussion.)

## DISCUSSION

**H.Rishbeth**

The author's formula $\sum = a\cos\chi + b\cos^{1/2}\chi + c$ may serve as an empirical formula, but I do not believe that it is a physical one. The $F$-layer contribution to the Pedersen conductivity would largely come from the $F1$-layer, for which the "$\cos\chi$ dependence" resembles that of the $E$-layer. Your results appeared to show a significant $F$-layer contribution to the Hall conductivity, which is difficult to understand.

**Author's Reply**

We agree that the given formula is an empirical one, although it is based on physical arguments. The fit to this formula by our data gives surprising results in the sense that $a > b$ for $\sum_H$. As our analyses show that the contribution to $\sum_H$ is negligible above 150 km, the results to the empirical fit is a contradiction as Rishbeth suggests. We are grateful for this comment and are looking into this problem.

**P.Vila**

Why does the ratio $\Sigma_H/\Sigma_P$ increase so much with precipitation?

**Author's Reply**

We are dealing with high-energy electrons, which populate only the 100–105 km levels.

# CALCULATION OF THE TEMPERATURE AND VELOCITY VECTOR OF THE NEUTRAL GAS USING EISCAT CP3-DATA

by

D. Figueroa and H. Kohl

Max-Planck-Institut für Aeronomie,
D-3411 Katlenburg-Lindau,
FRG

## ABSTRACT

The CP3-mode of EISCAT is particularly appropriate for three-dimensional studies of the interaction between ionosphere and thermosphere, because of its good latitude, height and time coverage. Using mainly CP3-data of EISCAT as input, we calculated parameters of the neutral gas in the middle thermosphere in the latitudes of EISCAT, by solving the coupled transport equations for the neutral and ion gases for the heights between 150 and 500 km and latitudes between 66° and 72°. By numerically solving the energy equation of the ions and the momentum equation of the neutral gas we obtained the temperature and the horizontal velocity of the neutral gas. This, used in combination with the energy equation of the neutral gas, allowed us to estimate the vertical neutral gas velocity.

Our calculated neutral gas temperatures sometimes deviate from those obtained by empirical models, which rather represent average conditions, and cannot be expected to described individual periods in the strong varying auroral thermosphere.

Our calculated neutral gas velocities show clearly the effect of the convection pattern of the ion gas, and little vertical variation. Our estimated vertical neutral gas velocities look fairly reasonable, although they are subject to several sources of error.

## INTRODUCTION

In middle latitudes the main energy input into the ionosphere is due to solar UV-radiation, which leads to ionisation of the medium and thereby to hot photoelectrons. Their energy randomizes and is transferred to the ions and from them to the neutral gas. In a state of equilibrium the ion temperature is, therefore, somewhat higher than the neutral gas temperature (a few to some hundred Kelvin depending on altitude) while the electrons usually have quite large temperatures (hundreds or a thousand Kelvin higher than the ions). It is, then, possible to determine the neutral gas temperature when ion and electron temperatures are known from measurements, e.g. from incoherent scatter observations (Carru [1]). In the polar ionosphere we have as additional heat sources penetrating energetic particles and especially strong electric fields that produce plasma drifts which finally dissipate into heat. It is particularly this latter effect which will be discussed in this paper.

## RESULTS

*Fig. 1 (left part)* shows ion temperatures $T_i$ measured by EISCAT for the period 8 L.T., 20/09/84 to 8 L.T., 21/09/84., the right part of the figure shows ion velocities $v_i$ for the same period. The measurements are done at an altitude of 325 km as a function of latitude and time. One realizes a strong relation between high ion velocity and high ion temperature, which is a consequence of the frictional heating between ions and neutral gas (Joule-heat).

The horizontal neutral gas wind may be estimated from the equation

$$\frac{\partial(v_i)_h}{\partial t} + 2(\Omega \times v_n)_h + \left(\mu \frac{\partial^2 v}{\partial z^2}\right)_h = -1/\rho(\Delta p_n)_h - \nu_{ni} \cdot (v_n - v_i)_h \qquad (1)$$

where the index h denotes the horizontal component. The second term on the left hand side is the Coriolis-force and the third term the kinematic viscosity. The driving force of the wind is on the right hand side: the horizontal pressure gradient, which was taken from the MSIS 86 model, and the ion drag force which is related to the measured ion velocity. The solution of this equation for the observational period of Fig. 1 is shown in Fig. 2. It is clearly seen that on the evening side the neutral wind follows the strong drag force exerted by the high ion velocities. On the morning side the neutral wind is mainly a result of thermospheric pressure gradients.

It is also possible to calculate the neutral temperature using the above results. The first method is to solve the energy balance equation of the ions, which is approximately

$$T_i = T_n + \frac{m_n}{3k}(v_i - v_n)^2 \qquad (2)$$

(e.g. Banks, [2]). As $T_i$ and $v_i$ are measured by EISCAT and $v_n$ is estimated as pointed out above, this equation can be solved for $T_n$. The results are shown as a function of time and latitude for 295 km altitude. It is clear from Fig. 3 that there is a strong temperature effect in the evening which is not related to solar UV-radiation, but is a consequence of Joule heat.

There is another way of calculating $T_n$, and that is by solving the energy balance equation of the neutral gas. This equation is very complicated and shall not be written down here explicitely. It contains besides the Joule heat term and the collisional energy transfer from hotter to colder particles, which are the same terms as in eq. (2), also heat input from UV-radiation or particles, thermal conduction, convection, radiative cooling etc. An attempt was made to solve this equation taking into account UV-input using a model of Banks and Kockarts [3], infrared cooling according to the same authors, thermal conduction, dissipation by viscosity, Joule heat and neglecting all other terms. The results are shown in Fig. 4. It turns out that there are serious discrepancies between both ways of calculation, the most obvious one is that temperatures calculated from the neutral energy equation are generally lower. The reason is probably that in the neutral gas energy balance equation some terms are neglected and others are not accurate enough.

An important contribution to neutral temperature variations can, for instance, come from vertical motions of the neutral air due to adiabatic cooling or heating. It was tried to attribute the difference between both the above solutions entirely to this effect. The necessary vertical motions, however, do not look very convincing; they are sometimes rather large (tens of m/s) and have an average downward velocity. A comparison of calculated $T_n$ and temperatures from atmospheric models as they are shown in Fig. 5 demonstrate that $T_n$ derived from the energy equation of the ions are not in good agreement. Deviations may be caused by local or temporal effects, that, of course, cannot be accounted for by global models

The collisional model used so far in this paper includes polarization and resonant interaction between ions and neutrals (Banks, [4]) and was corrected to take into account the effects of the velocity distribution using the hard-sphere-model (Schunk, [5]). In order to give a better approximation a more sophisticated model (Figueroa and Hernandez, [6]) was used instead of the hard-sphere-model. The differences in the derived neutral gas velocities and temperatures are significant (15 %, respectively 5 %), both parameters becoming larger.

## REFERENCES


1 Carru, H., M. Petit and P. Waldteufel, On the diurnal variation of the thermopause temperature, Planet. Space Sci., **15**, 944-945, 1967.

2 Banks, P.M., in "Exploration of the upper atmosphere", edited by C.S. Deehr and J.A. Holtet, pp. 113-127, D. Reidel, Dordrecht/Boston 1981.

3 Banks, P.M. and G. Kockarts, in "*Aeronomy*", *Academic Press*, 1973.

4 Banks, P.M., Collison frequencies and energy transfer, Planet. Space, Sci., **14**, 1105-1122, 1966.

5 Schunk, R.W., Transport equations for aeronomy, Planet. Space Sci., **23**, 437-485, 1975

6 Figueroa, D., and R. Hernandez, Collisional transfer rates for momentum and energy exchange in the case of relative drifts between ions and their parent neutral gas, J. Atmosph. Terr. Phys., **50**, 447-454, 1988.

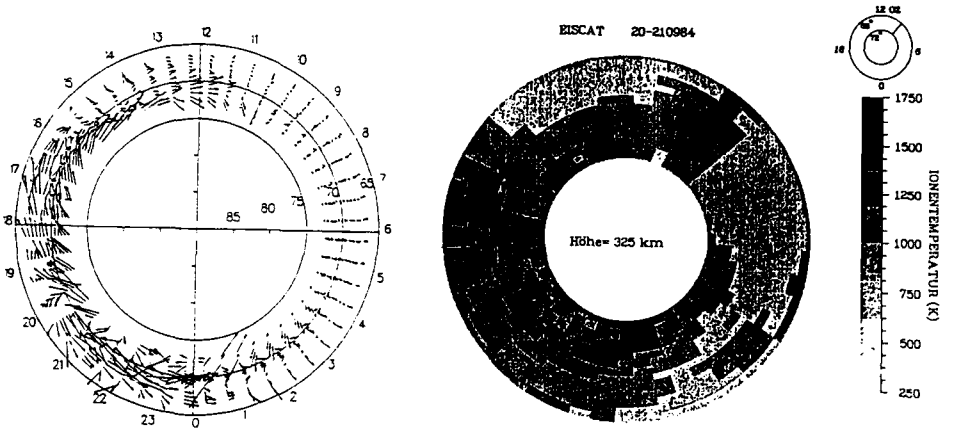



Figure 1. Ion temperatures measured by EISCAT at 325 km altitude as a function of local time and latitude (left). Polar diagram of the ion velocity vector for the same period (right). Date: 20./21.09.84.

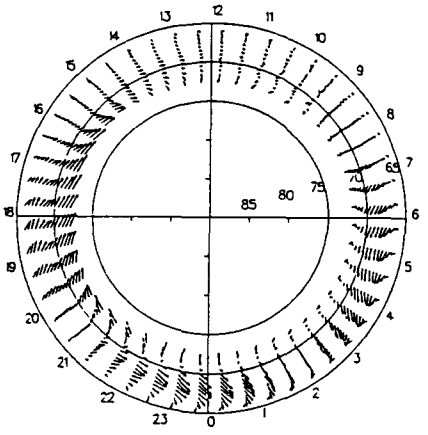


Figure 2. Calculated neutral wind derived from the ion velocity measurements.

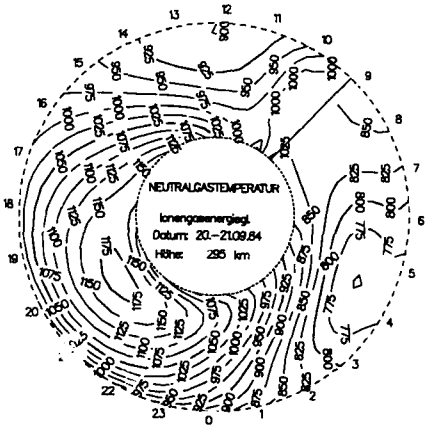



Figure 3. Neutral gas temperature obtained from the ion energy balance equation (3).

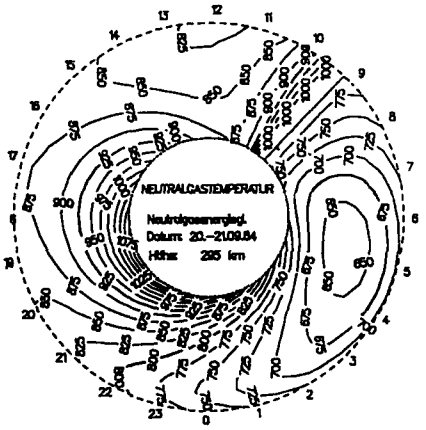



Figure 4. Neutral gas temperature obtained from the neutral gas energy balance equation.

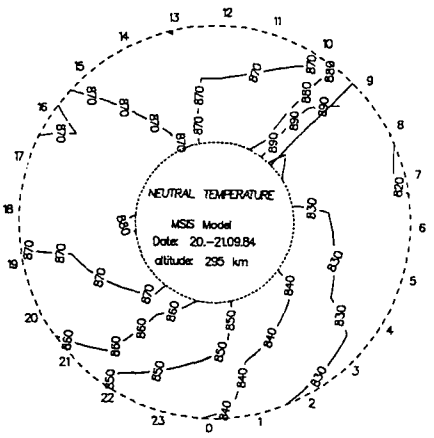



Figure 5. Atmospheric model temperatures (MSIS) for the same time.

## DISCUSSION

**K.Rawer**

*After* your tests, what would be your estimate of the uncertainty of the neutral temperature and velocity data routinely published by a few incoherent scatter stations?

**Author's Reply**

I do not know these data, probably for middle latitudes. I made my calculations for high latitudes. It is difficult to evaluate the errors of my results, because these depend on several models.

**D.Rees**

The derivation of neutral wind and temperature from radar observations has been an attractive possibility for many years. However, of three possible parameters, only meridional winds have yet been demonstrated to be *always* consistent with real thermospheric neutral motions, as seen by satellite sensors or Fabry-Perot interferometers. To proceed further than meridional winds, great care or infinite luck is required!

Firstly, the radar does not measure vertical neutral wind, which is thus a significant error source in deriving meridional wind from parallel ion drift. Secondly, correction for at least a bi-Maxwellian ion velocity distribution is essential. Thirdly, the real atmosphere will depart from any model atmosphere ($\Delta T$; $\Delta\rho$; $\Delta M$ (mean mol. mass)). Since these are compounding, non correlated errors in the radar analysis, the neutral meridional wind may well be in error by 200 m $s^{-1}$ or $\pm 50\%$ (take the larger!)

To anaiyze temperature ($T_N$) and then zonal winds, the above error on meridional wind, and model errors, compound. Under some ideal conditions, the zonal wind, and $T_N$, *may* be valid. However, most of the time, the compounded effects of errors, and lack of knowledge of *global* changes of the thermosphere, will create major errors, but without providing a reliable "flag" that an error is present.

I would thus stress very great caution when interpreting the present results, and recommend intercomparison with direct wind measurements by Fabry-Perot interferometers, etc.

**Author's Reply**

No response to comment.

# ION HEATING EVENTS OBSERVED BY THE EISCAT RADAR

**I.W. McCrea, T.R. Robinson, M. Lester, & T.B. Jones.**

Ionospheric Physics Group
Department of Physics
University of Leicester
LEICESTER LE1 7RH
U.K.

## Abstract

According to simple theory, a quadratic relationship should exist between ion temperature and ion velocity during frictional heating of the high-latitude ionosphere. A method of fitting the form of this relationship for a selection of EISCAT Common Programme data is demonstrated, and the observed behaviour is consistent with theoretical predictions. From the equation obtained, estimates are made of the neutral mass and one component of the neutral wind. These appear to be physically realistic. The validity and sensitivity of the technique is briefly discussed, and an attempt is made to account for possible errors due to changes in the ion composition and the presence of anisotropic plasma velocity distributions.

## Introduction

It is well established that, in high-latitude regions, frictional or 'Joule' heating makes a major contribution to the atmospheric energy balance. Such heating arises when magnetospheric electric fields drive the ion population through the neutral atmosphere at high velocity. Collisions between the ions and neutrals convert the energy from directed motion into heat. *Banks*, [1], and *Banks and Kockarts*, [2] have obtained the full form of the ion energy equation which includes a treatment of the effects of conduction, convection and energy exchange between particles. Various workers ( e.g. *Schunk and Sojka*, [3], *Conrad and Schunk*, [4], and *St. Maurice and Hanson*, [5] ) have demonstrated that the full equation can be substantially simplified for a realistic ionosphere and can be expressed in the form:

$$\sum_n \frac{n_i \upsilon_{in} \mu_{in}^2 ( \underline{V}_i - \underline{V}_n )^2}{m_i} + 3n_i \upsilon_{ie} k ( T_e - T_i ) = \sum_n \frac{3\mu_{in}^2 n_i \upsilon_{in} k ( T_n - T_i )}{m_i m_n} \qquad (1.1)$$

Where:

| | | | |
|---|---|---|---|
| $n_i$ | is the ion density. | $m_i$ | is the ion mass |
| $\underline{V}_i$, $\underline{V}_n$ | are the ion and neutral velocities. | $m_n$ | is the neutral mass |
| $k$ | is the Boltzmann constant | $T_i$ | is the ion temperature |
| $T_n$ | is the neutral temperature. | $T_e$ | is the electron temperature |
| $\mu_{in}$ | is the reduced mass, given by $(m_i m_n) / (m_i + m_n)$ | $\upsilon_{in}$ | is the momentum transfer collision frequency. |

Note that the energy sources and sinks for the ion gas have been separated to opposite sides of the equality. Ions may gain energy by interactions with electrons, or acceleration by applied fields, and this energy is dissipated by collisions with the neutral air.

A further simplification of this equation [5], is of the form:-

$$( \underline{V}_i - \underline{V}_n )^2 = \frac{3k}{m_n} ( T_i - T_{eq} ) \qquad (1.2)$$

where $T_{eq}$ is the sum of the neutral temperature and the electron energy exchange term.

From this, one obtains by simple vector manipulation,

$$T_i = \frac{m_n}{3k} V_i^2 - \frac{2m_n V_n \cos ( A )}{3k} V_i + \frac{m_n V_n^2}{3k} + T_{eq} \qquad (1.3)$$

where A is the angle between the ion and neutral flows.

This simple theory predicts a quadratic relationship between the ion temperature and ion velocity. The simplified expression (1.2) has formed the basis of several attempts to study the coupling between the ionised and neutral atmospheres. Neutral temperature and velocity have been inferred by means of ion temperatures and velocities measured with the Atmospheric Explorer C satellite [5] based on some assumptions regarding the spatial variation of $T_n$ and $V_n$. *Baron and Wand,* [6], observed ion temperature enhancements due to frictional heating with the incoherent scatter radar at Chatanika, Alaska. They reported that an asymmetry existed between the ion temperatures measured in the eastward and westward electrojets for a given magnitude of plasma velocity. *Alcayde et al,* [7], attributed this effect to the influence of the Coriolis force on ion-neutral coupling. *Lathuillere and Brekke,* [8], have speculated that variations of ion composition may also be involved. *Alcayde and Fontanari,* [9], attempted to derive the neutral wind vector from incoherent scatter data, by first deducing an empirical model of the neutral temperature. *Killeen et al,* [10], attempted to evaluate each term of the theoretical expression with the aid of simultaneous measurements of the ion and neutral temperatures and velocities obtained from instruments aboard the Dynamics Explorer satellite. In the majority of cases, the ion energy balance equation was found to be adequate to describe the variations observed.

In this paper, the form of the relationship is tested by means of data taken with the European Incoherent Scatter Radar (EISCAT), and the equation obtained is interpreted in terms of the neutral mass and neutral velocity. In the present study an attempt is made to dispense with the need to assume model parameters when obtaining the coefficients of the ion energy equation. This is achieved by simultaneous fitting of the three coefficients of the quadratic. In this approach, some reduction in the range of parameters and in the time resolution must occur.

EISCAT is a tristatic radar system located in Northern Scandinavia. Details concerning its construction and capabilities have been extensively reviewed by previous authors (e.g.*Rishbeth and Williams,* [11], *Folkestad et al,* [12] ). The data presented here were obtained from operation of the radar in the C-P-1 mode, in which the Tromsø transmitter beam is fixed parallel to the local magnetic field line, and the remote receivers scan this beam, intersecting it at heights of 110, 119, 131 and 312 km. The availability of tristatic measurements enables the full vector of the ion velocity to be resolved, a great advantage over previous monostatic systems, which relied on beam-swinging techniques. Ionospheric parameters are obtained by fitting the measured autocorrelation function to one of a range of theoretical functions, corresponding to a known set of parameter values. This theoretical A.C.F. is continuously corrected, by means of a non-linear, least-squares Newton-Raphson technique, until a satisfactory fit is obtained.

### Fitting to the Ion Energy Equation

Data from four days have been chosen from the extensive EISCAT CP-1 data set, according to the following criteria:-

(i) The data contained evidence of at least one large ( >200K ) perturbation of the F-region ion temperature, corresponding to an enhancement of the measured ion velocity. The value of 200K was chosen since this is the largest temperature enhancement that could be caused by electron-ion energy exchange effects alone [5]. In these cases, the correlation between the temperature and velocity enhancements provides convincing evidence that the observed heating was due to the action of frictional effects.

(ii) Data were recorded continuously throughout the experiment at all three sites.

(iii) The data were selected from relatively recent experiments, since these should yield higher quality data than experiments performed in the early days of EISCAT.

The dates and times of the selected experiments are summarised in Table 1, along with the $K_p$ sum over a given number of three hour epochs, the number of three hour epochs included in the summation, and the average activity over the period. The experiments in general cover periods of moderate activity, while at times activity is low. The two epochs of highest $K_p$, when the index reached values of $6^0$ and $5^0$ respectively, occurred on the nights 12-13 / 11 /1985 and 25-26 / 3 / 1986.

**Table 1**

| Start Date and Time | End Date and Time | Kpsum | Epochs | Average $K_p$ |
|---|---|---|---|---|
| 10.09.1985 08:00 | 11.09.1985 22:00 | $29^+$ | 13 | $2^+$ |
| 12.11.1985 09:00 | 13.11.1985 23:00 | $32^+$ | 13 | $2^+$ |
| 25.03.1986 09:00 | 26.03.1986 09:00 | $27^0$ | 8 | $3^+$ |
| 08.04.1986 08:00 | 09.04.1986 08:00 | $16^-$ | 8 | $2^0$ |

The variation of the ion temperature with time during the chosen intervals at a height of 312 km.is reproduced in Figure 1. In the data fitting, the ion composition is assumed to be dominated by $O^+$. The variation of the ion velocity magnitude derived from tristatic measurements at the same height is also included in this figure. The small 'H' at the top of the graph marks the approximate location of the Harang discontinuity, where the zonal plasma velocity changed from westward to eastward. The exact time of the discontinuity on 8th April, 1986 proved difficult to estimate, due to the low velocities in this region.

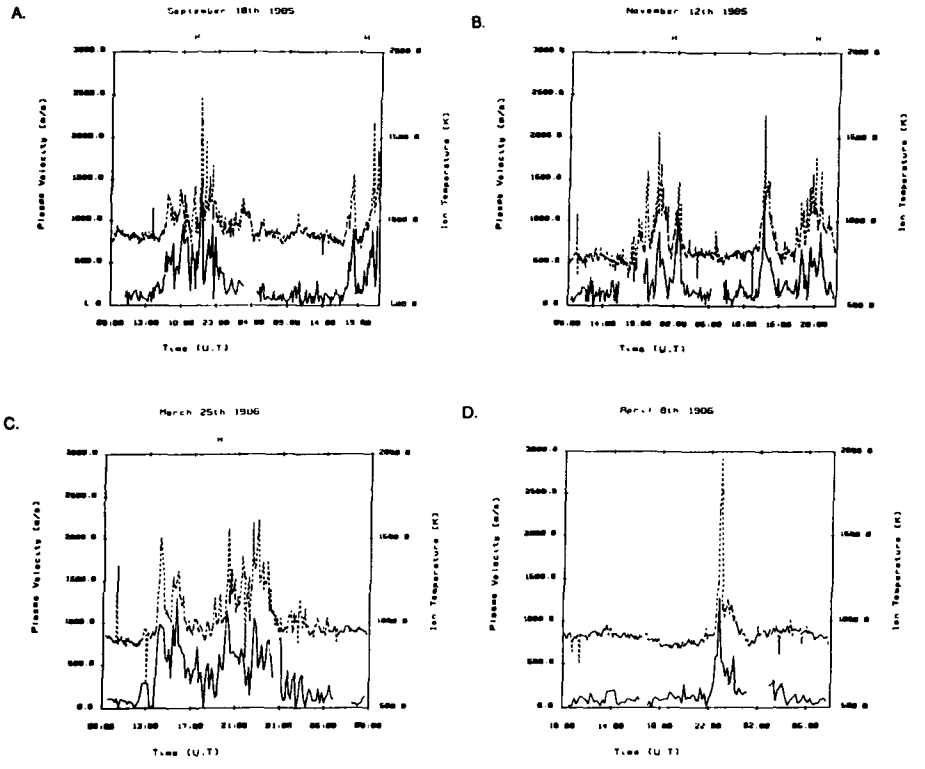



**Figure 1.** Variation of ion temperature (dotted line) and ion velocity (full line) for the four days of the study. The small 'H' marks the position of the Harang discontinuity.

Clear correlation of the ion temperature enhancements with large ion velocities can be seen on all four days. Previous workers ( e.g. [5] ) have commented that a close correlation of $T_i$ with $V_i$ indicates weak coupling between the ion and neutral velocities, since in such a case the ion-neutral difference velocity would be largely dependent on fluctuations in $V_i$.

The average ion temperature is illustrated in Figure 2 as a function of the total plasma velocity, with the sign convention chosen so that eastward (westward) directed velocities are labelled positive (negative). Three of the intervals displayed significant ion temperature enhancements before and after the Harang discontinuity. Although there is a degree of asymmetry in the temperatures associated with the two electrojets, the differences shown in Figure 2 are not as large as those found by *Baron and Wand,* [6]. However, the non-linear form of the relationship between $T_i$ and $V_i$ is clearly evident.

In order to ascertain the form of this relationship, an attempt was made to fit the temperature response with polynomials of varying degree. A least-squares fitting technique was employed, and fits were attempted for linear, quadratic, cubic and quartic equations. In Figure 3, height profiles of the correlation coefficients obtained by fitting to a linear variation of $T_i$ with $V_i$, and a quadratic in $V_i$ are reproduced. *Rishbeth and Ashford,* [13], have attempted to fit a linear relationship to the variation of $V_i^2$ with $T_i$ during a run of the UK-POLA experiment, the implicit assumption being that the mean neutral wind magnitude is close to zero. Such a fit has also been attempted for the present data, and the correlation coefficient profile obtained by this method is also depicted in Figure 3. It is clear that the quadratic relationships represent significantly better agreements than the linear variation of $V_i$ with $T_i$. High correlation coefficients were found for equations of order higher than two, but in these cases the coefficients for the higher powers were always observed to be extremely small, ( $< 10^{-6}$ ).

The fitted equations were found to have the smallest correlation coefficients at heights around 200km. In this height range, the neutral and ion compositions are known to be changing rapidly with altitude, and it is possible that composition changes could occur on time scales comparable with the data lengths to which the fitting has been performed. At F-region heights, the correlation coefficient for the quadratic form of the equation is of order 0.9. The form of this variation is in accord with the simple theory presented in the previous section.

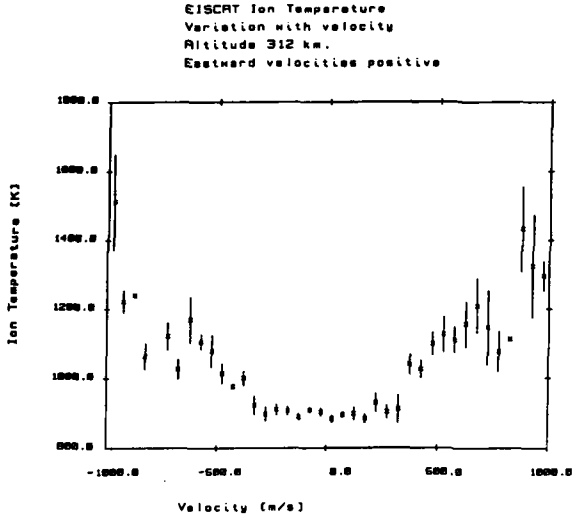



**Figure 2**

Variation of ion temperature with ion velocity magnitude, with eastward flow denoted positive and westward flow negative. The points shown were deduced by binning the complete four day data set. The non-linear nature of the response is clearly shown.

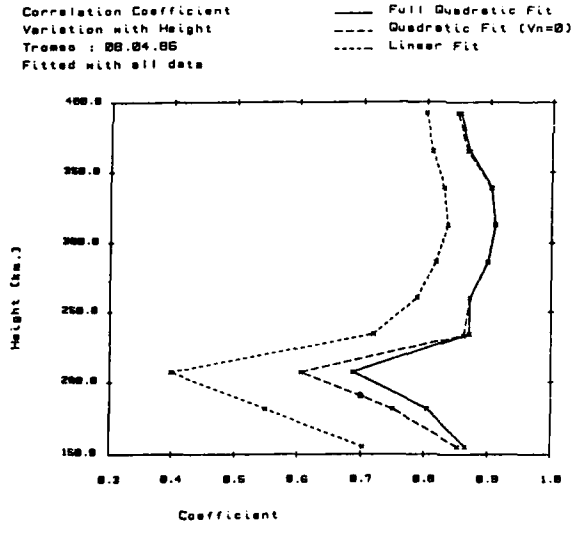



**Figure 3.**

Profiles of correlation coefficient obtained by fitting the relationship between ion temperature and velocity to linear and quadratic forms. Note that the largest coefficients are obtained in the case of the quadratic fit, and that the fit is optimal at F–region heights.

In the determination of the correlation coefficients ( Figure 3 ), the data series was several hours in length, with as many points as possible being taken in order to define the nature of the relationship. To develop this method further, it was necessary to demonstrate that the quadratic nature of the equation relating $T_i$ and $V_i$ was still evident during much shorter data lengths. The magnitude of the correlation coefficient was investigated for various data lengths. The resolution of the data points was five minutes, so that a ten point series corresponded to a running average over fifty minutes. Data series with lengths ranging from five to sixteen points, were investigated. Correlation coefficients in excess of 0.9 were obtained even for the shortest data series, provided the series embraced a reasonable range of ion temperature and velocity. For series during which no large variation in these parameters occurred, the correlation coefficient was invariably lower and large variations were observed in the coefficients.

**Determination of Geophysical Parameters**

The validity of the parameters derived from the fitted coefficients of the ion energy equation was determined by calculating the neutral mass over a range of altitudes. *Rishbeth*, [14], has pointed out that it should be possible to determine the neutral mass from the form of the ion energy equation (1.3), since the neutral mass should be directly proportional to the coefficient of the squared term. From the derived polynomials, the quadratic coefficient is of the correct order of magnitude to yield values of the neutral mass comparable with model predictions. In Figure 4, height profiles of the neutral mass obtained from the two quadratic fits are compared with the height variation predicted by the MSIS86 model for a period encompassing the intense heating event of April 8th 1986. The agreement is within 15% in the F-region, but becomes poor in the transition region, possibly due to the variation of composition described earlier. During the fitted period, the Joule heating event would be expected to cause an upwelling of molecule-rich air from E-region heights. Hence it is not unreasonable to assume that a variation in the neutral mass could occur in the lower F-region at these times.

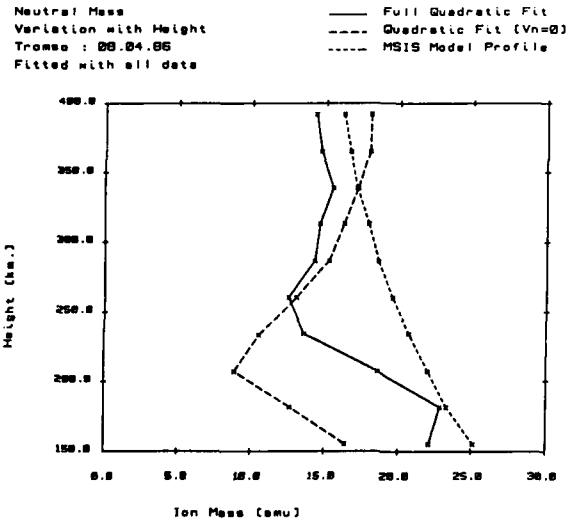



**Figure 4.**

Profiles of the neutral mass, derived by assuming :-
(i) A non-zero neutral wind,
(ii) A neutral wind which averages to zero.

These profiles are compared with the MSIS86 profile. Note that in the F-region, the experiment and theory are in good accord.

It may be noted that the quadratic corresponding to the case of no neutral wind (where the linear term of the equation has been set to zero) produces slightly better fits to the model profile in the F-region than the full form of equation (1.3). Given the small magnitudes of the coefficients involved, however, together with the possibility of inadequacy in the neutral composition model, it is impossible to establish whether such a simplification produces an improvement upon (1.3).

The validity of the linear term was investigated by considering its physical significance. From (1.3), the linear coefficient is found to be proportional to $V_n CosA$, which corresponds to the component of the neutral velocity in the direction of the plasma drift, since A is the angle between the ion and neutral flows. By adopting a value of 21 a.m.u. for the F-region neutral mass, the magnitude of this component was obtained from the fitted linear coefficient.

Various numbers of points were tried in this fit. With data taken at five minute resolution, a five point series represents a running average over 25 minutes. Stepping this series sequentially by one point across the heating period gave consistent estimates of this velocity, although some large values were obtained toward the end of the event. The values for the velocity component are within the range of previous measurements ( see *Straus*, [15], for a review ). Figure 5 represents the variation of the velocity component as a function of time for different data lengths. While the consistency improves as the data length increases, there is a corresponding decrease in time resolution. A data series of some tens of minutes should however be sufficient, since the ion-neutral coupling time constant is predicted to be around three hours for this event, and is of the form [6] :-

$$\tau = \frac{0.226 \times 10^6}{N_i} \qquad (2.1)$$

The long time constant is thus due to the low ambient electron density at the time of this heating event.

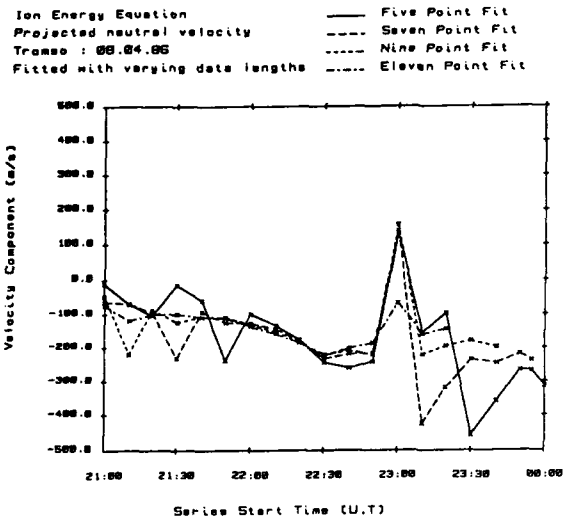



**Figure 5.**

Component of the neutral wind in the direction of the ion flow, derived from the linear coefficient of the ion energy equation, fitted with data series of different lengths. The consistency of the estimates improves with the use of longer series, although this is achieved at the expense of time resolution.

An attempt was then made to verify the values obtained for this velocity component. The method followed the procedure of *Amayenc and Vasseur*, [16], who reported measurements of the meridional component of the neutral wind, derived from field-aligned velocities after allowing for the influence of the ion diffusion velocity. The full formula relating parallel velocity to neutral wind is of the form:-

$$V_{para} = V_{nm} \cos(I) - \frac{2kT_i \sin(I)}{m_i}\left[\frac{(T_e+T_i)}{2T_i}\cdot\frac{\partial N_e}{\partial z} + \frac{1}{2T_i}\cdot\frac{\partial(T_e+T_i)}{\partial z} + \frac{m_i g}{2kT_i}\right] \qquad (2.2)$$

Where: $V_{nm}$ is the meridional component of the neutral wind
$I$ is the inclination angle of the geomagnetic field
with the other variables having been defined previously.

Although this method gave realistic values for the meridional wind component, checking of the component parallel to the plasma drift required knowledge of the zonal wind vector also. An attempt to solve the simplified ion energy equation for the neutral velocity magnitude by means of an empirical model of the neutral temperature proved unsuccessful, due to some large variations in the apparent direction of the plasma flow.

## Discussion

The technique of fitting curves to the variation of ion temperature as a function of ion velocity relies on obtaining accurate measurements of both of these parameters. The method of assigning errors to the measured EISCAT velocity has been reviewed by *Jones et al*, [17], where it is noted that the major source of error is due to the small value of the signal to noise ratio, particularly at the remote sites. In multipoint fitting, random errors of this nature should be averaged out by the inclusion of a large number of points. The accurate measurement of the ion temperature poses more serious problems, since determination of the true value is likely to be affected by both systematic and random errors.

At least two mechanisms could give rise to systematic errors in the determination of the ion temperature. Firstly, all of the plasma temperatures presented up to this point have been calculated from the recorded A.C.F. by means of a standard model for the ion composition. This does not include processes such as atmospheric heating although it is known that these effects can radically alter the ion composition, even at F-region heights ( [18] ). Composition changes might be expected for the conditions of interest here, and some allowance must be made for the possible underestimate of $T_i$ that might occur if too low a mass had been adopted in the fitting of the $T_i / m_i$ ratio. The fitting procedure was thus modified in order to allow for this possibility.

Various workers, ( e.g. *Lathuillere et al*, [19], *Lathuillere and Brekke*, [8] ) have adopted a technique which consists of fitting once with the modelled composition, and then making a second fit to the A.C.F, starting from the results of the first fit, but with the ion mass included as an additional free parameter. This approach is prone to inaccuracy, since the effect of changing composition upon the autocorrelation function is small. Nonetheless, this technique was attempted in addition to the default method in order to obtain a qualitative picture of variations in the composition.

In Figure 6, values of the ion mass profile from the default analysis model are compared with those obtained by fitting to the ion mass, during a period without heating, and from a period where intense heating was occurring. Points which converged to an acceptable fit are marked with a cross, those which did not are marked with an open circle. Although the default model is in broad agreement with the profile obtained at low heating rate, the F-region ion mass is noticeably increased over the normal values during.the period of high heating.

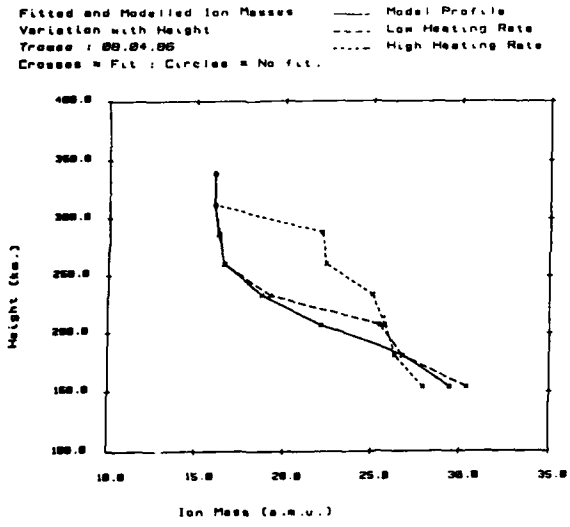



**Figure 6.**

Fitted ion composition profile, compared to the EISCAT default model, during periods of low and high frictional heating. Note the increase in mass during the strong event, possibly due to the upwelling of molecule-rich neutral air from the heated E-region.

The effect of this on the fitted temperatures is seen in Figure 7, where the ion temperature profiles derived from the default compositions during the times of low and high heating are compared with those derived from the fitted compositions. If one disregards the point at 207 km., for which no fit was actually achieved, the profiles at the time of low heating are in good accord. However, the temperatures during periods of high heating appear to be seriously underestimated by the default model, over a large range of heights. This effect chiefly appears in the bottomside F-region. At greater heights, fitting to the composition is difficult, but the results indicate that the perturbation is not of great significance, even in the most intense of the observed events. It therefore seems probable that calculating ion temperatures on the basis of an $O^+$ ionosphere gives adequate values of ion temperature for altitudes above the F-region peak, although in the bottomside F-region, the default composition is clearly inadequate at times of high heating. Such results for the variation of temperature with ion composition agree with the trends noted in a recent paper by *Kofmann and Lathuillere*, [20]. The method of performing two fits is however not consistently successful, and the results cannot be employed in the same way as those obtained by means of the default model.

An additional problem in measuring the ion temperature is that the measured temperature may not be appropriate to the energy equation. This can occur if an anisotropic plasma velocity distribution exists in the observing volume, in which case the measured temperature is dependent on the particular velocity component observed. In recent papers, ( [21], [22], [23], [24] ) the existence of anisotropic distributions in the auroral ionosphere has been widely reported. These distributions take the form of toroids in velocity space, with the highest plasma velocities encountered perpendicular to the magnetic field. A detailed discussion of these effects is beyond the scope of the present report (see *St. Maurice and Schunk*, [21] for a full review), but they have a profound consequence for any measurement of temperature during periods of high velocity. In these circumstances,the measured temperature will also be anisotropic, with the temperature parallel to the field being an underestimate of the 'thermodynamic' temperature required by the energy balance equation (e.g *Perrault et al*, [22], *Lovhaug and Fla*, [23] ).

Although in the present case, it seems that the ion-neutral difference velocity will only slightly exceed 1 km/s, at which point the error in the estimate of the temperature caused by the anisotropy begins to exceed the inherent error in the measured temperature, the relevant corrections have been included here for completeness. The formulae employed are due to *St. Maurice and Schunk*, [21], [25], with correction terms given by *Perrault et al*, [22], and adopted subsequently by *Williams and Jain*, [26].

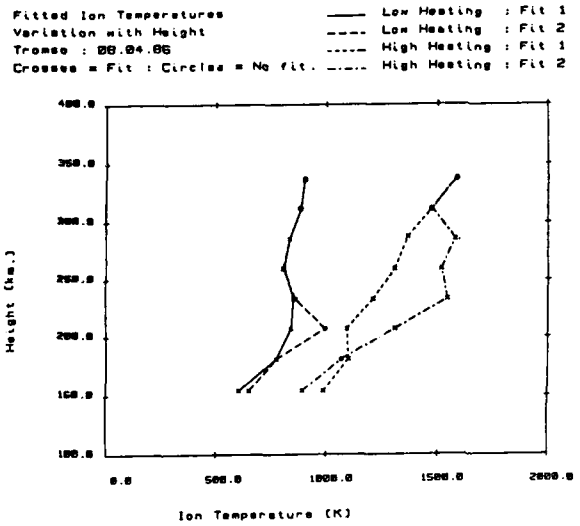



**Figure 7.**

Temperature profiles fitted with the default model, and the derived compositions for low and high heating rates. Note that the default model appears to underestimate the temperature during the strongest heating events.

The corrected temperature is within the errors of the fitted ion temperature at Tromsø throughout the event, the single exception being at the peak of the heating, where the difference velocity reached some 1200 m/s. At this time, the corrected temperature was around 30% in excess of the parallel ion temperature. Although anisotropy errors seem unlikely to be a major influence in the present data, these results suggest that they should be of considerable importance for any event in which large difference velocities are maintained for a long period of time.

If the correct plasma temperature and velocity can be accurately measured, consideration should be also given to the validity of the ion energy equation itself. *Killeen et al*, [10], were able to evaluate both sides of this equation, by means of simultaneous measurement of the temperature and velocity of both the ionised and neutral components of the atmosphere. Their investigation revealed that although the measured ion temperature was generally in good agreement with that predicted from the calculated ion heating rate, there were periods when significant discrepancies were observed. These were attributed to the action of an additional heat source for the ions, not accounted for by the simple theory. This mechanism is still an unsolved problem, and requires further studies involving simultaneous measurement of ion and neutral parameters.

## Summary and Conclusions

Data have been presented from four runs of the EISCAT Common Programme, CP-1, for which the plasma density, velocity and temperatures have been derived. By employing a least-squares polynomial fitting procedure, the variation of ion temperature with plasma velocity can be investigated, giving the quadratic form predicted by simple theory. The form of the polynomial has been inspected in an attempt to infer information on neutral mass and a component of the neutral velocity. These procedures yield acceptable values for the coefficients of the energy equation, however great caution must be exercised in application of these methods to quantitative studies, since there are a variety of potential errors, which can be maximised during the most geophysically interesting periods.

The influence of factors such as ion composition and anisotropic velocity distributions cannot be neglected in studies of this kind. Attempts at eliminating these effects by adopting a more sophisticated analysis have been described. The simple approximation to the energy equation seems to be valid for measurements in the high F-region, provided that the ion-neutral difference velocity is less than approximately 1 km/s.

**List of references**


| | | | | | |
|---|---|---|---|---|---|
| 1 | Banks | Planet.Space Sci. | 14 | 1105 | 1966 |
| 2 | Banks and Kockarts | 'Aeronomy', Academic Press, New York | | | 1973 |
| 3 | Schunk and Sojka | J.Geophys.Res. | 87 | 5159 | 1982 |
| 4 | Conrad and Schunk | J.Geophys.Res. | 84 | 811 | 1979 |
| 5 | St.Maurice and Hanson | J.Geophys.Res. | 87 | 7580 | 1982 |
| 6 | Baron and Wand | J.Geophys.Res. | 88 | 4114 | 1983 |
| 7 | Alcayde et al | J.Atmos.Terr.Phys. | 46 | 625 | 1984 |
| 8 | Lathuillere and Brekke | Ann.Geophys. | 3 | 557 | 1985 |
| 9 | Alcayde and Fontanari | J.Atmos.Terr.Phys. | 48 | 931 | 1986 |
| 10 | Killeen et al | J.Geophys.Res. | 89 | 7495 | 1984 |
| 11 | Rishbeth and Williams | Quart.Journal R.A.S | 26 | 478 | 1985 |
| 12 | Folkestad et al | Radio Sci. | 18 | 867 | 1983 |
| 13 | Rishbeth and Ashford | to be submitted (Nature) | | | 1988 |
| 14 | Rishbeth | Private communication | | | 1987 |
| 15 | Straus | Rev.Geophys.Sp.Phys. | 16 | 183 | 1978 |
| 16 | Amayenc and Vasseur | J.Atmos.Terr.Phys. | 34 | 351 | 1972 |
| 17 | Jones et al | J.Atmos.Terr.Phys. | 48 | 887 | 1986 |
| 18 | Schunk et al | J.Geophys.Res. | 80 | 3121 | 1975 |
| 19 | Lathuillere et al | Radio Sci. | 18 | 887 | 1983 |
| 20 | Koffmann and Lathuillere | Geophys.Res.Lett. | 14 | 1158 | 1987 |
| 21 | St.Maurice and Schunk | Rev.Geophys.Sp.Phys. | 17 | 99 | 1979 |
| 22 | Perrault et al | J.Atmos.Terr.Phys. | 46 | 531 | 1984 |
| 23 | Lovhaug and Fla | J.Atmos.Terr.Phys. | 48 | 959 | 1986 |
| 24 | Lockwood et al | to be submitted (J.Geophys.Res.) | | | 1988 |
| 25 | St.Maurice and Schunk | Planet.Space Sci. | 25 | 243 | 1977 |
| 26 | Williams and Jain | J.Atmos.Terr.Phys. | 48 | 423 | 1986 |



**Acknowledgements**

The authors would like to acknowledge the efforts of the EISCAT staff in their operation of the facility, and the staff of the Rutherford Appleton Laboratory for maintaining the data analysis software. Special thanks are due to Nigel Wade for his assistance at all stages. Part of this study was undertaken while one of us (IWM) was a guest at the Max-Planck Institut fur Aeronomie, Lindau, FRG. He would thus like to thank the M.P.I. staff, and particularly Dr.K.Schlegel, for some most helpful ideas and discussions.

EISCAT is an international facility, funded by Suomen Akatemia (Finland), the Centre National de la Recherche Scientifique (France), the Max-Planck Gesellschaft (West Germany), the Norges Almenvitskaplige Forskninsgrad (Norway), the Naturvetenskaplig Forskningsradet (Sweden), and the Science and Engineering Research Council (United Kingdom).

## DISCUSSION

**K.J.Winser**

Mr. McCrea indicated that a possible source of error in his interpretation of the data obtained in the lower $F$-region could be due to the ion composition model assumed in the analysis program. This is a very important point. Changing the model in this region by only 30% or so can affect the derived ion and electron temperatures by several hundred degrees. It should be noted also that most incoherent scatter radar facilities assume a static ion composition model in their data analysis programs. These models do not change for season, or different geophysical conditions, and therefore constitute a major limitation on ground-based measurements of the lower $F$-region.

**Author's Reply**

I agree that his was certainly the case, and that this point is essential to the interpretation of fitted temperatures. Some attempt was made in the paper to address this point by employing a multiple-fitting technique. Such procedures are difficult up the field- ine, since the velocity distribution function is not Ramon-like at these aspect angles, so that the distributions for various species are not well differentiated. Errors in the composition can certainly account for the disparity between the fitted neutral mass value and the model prediction in the lower $F$-region, where the composition is changing rapidly.

# ON THE INFLUENCE OF IONOSPHERIC CONDUCTIVITY ON DISPERSION OF ACOUSTIC GRAVITY WAVES

L.P.J. Kamp, R.H.M. Miesen, P.C. de Jagher, F.W. Sluijter
Department of Physics, Eindhoven University of Technology,
5600 MB Eindhoven, The Netherlands

**Abstract**

The influence of ionisation and conduction on the dispersion of acoustic gravity waves in the upper atmosphere has been investigated by means of a numerical analysis of the corresponding dispersion relation in which the Hall frequency is assumed to be much smaller than the Brunt-Väisälä frequency. It is shown that a significant consequence of the incorporation of ionospheric conduction is a relatively strong damping of gravity waves, unless they propagate parallel to the earth's magnetic field with frequencies smaller than the Brunt-Väisälä frequency. Typically this damping takes place with a characteristic damping rate that scales with the Pedersen conductivity. Furthermore it turns out that waves that belong to the acoustic branch of the double-branch dispersion relation are also damped with a damping rate of the same order in magnitude when ionospheric conduction is taken into account. This is the case for waves propagating perpendicular to the earth's magnetic field and for waves propagating parallel to the earth's magnetic field but then with frequencies that are of the order of the acoustic cut-off frequency.

## 1. Introduction

As is well-known, due to the stratification of the earth's atmosphere, there can propagate so-called internal acoustic gravity waves that have frequencies of the order of the Brunt-Väisälä frequency (Hines, 1960).

At ionospheric heights in the atmosphere a non-zero conductivity causes a coupling between these acoustic gravity waves and electromagnetic oscillations. This coupling results from collisions between charged and uncharged particles and makes it difficult to examine the dispersion properties of the acoustic gravity waves.

In this paper we will show that for atmospheric values of the relevant parameters the electric field is not negligible but most of the time hydromagnetic effects can be accounted for by incorporating only a term that represents the Lorentz force in the momentum equation. This is demonstrated by comparing the local dispersion curves for hydromagnetic-electromagnetic waves with dispersion curves of the dispersion relation for acoustic gravity waves taking only the Lorentz force into account. The main effect of ionisation will be damping of the acoustic gravity waves.

In special cases, when the Hall conductivity becomes large enough so that the Hall frequency is of the same order as the Brunt-Väisälä frequency, the properties of acoustic gravity waves can be affected by the induction of Hall-currents. In that case also the real parts of the frequency and the wave-number are affected significantly.

## 2. The dispersion equation

We can describe the hydrodynamic and electromagnetic waves in the atmosphere by using the following equations (Hines, 1953)

$$\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho \mathbf{u}) = 0, \tag{1}$$

$$\rho\frac{D\mathbf{u}}{Dt} = -\nabla p + \rho\mathbf{g} + \mathbf{J}\times\mathbf{B}, \tag{2}$$

$$\frac{D}{Dt}\left[p\rho^{-\gamma}\right] = 0, \tag{3}$$

$$\mathbf{J} = \vec{\sigma}\left[\mathbf{E} + \mathbf{u}\times\mathbf{B}\right], \tag{4}$$

$$\nabla\times\mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{5}$$

$$\nabla\times\mathbf{B} = \mu_o\mathbf{J} + \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t}. \tag{6}$$

The first equation is the continuity equation where $\rho$ is the mass density and $\mathbf{u} = (u,v,w)$ is the velocity of the air. Equation (2) is the momentum equation with the Lorentz-force $\mathbf{J}\times\mathbf{B}$ the pressure gradient $\nabla p$, and the gravitational acceleration $\mathbf{g} = -g\mathbf{1}_z$ balancing the inertia force. The equation of state (3) where $\gamma$ is the ratio of specific heats, is the one for isentropic processes which can be shown to be valid for acoustic gravity waves up to about 300 km. Above 300 km heat conductivity can not be neglected anymore. The current density $\mathbf{J}$ is given by a generalized Ohm's-law (4) where the conductivity tensor $\vec{\sigma}$ has non-vanishing elements $\sigma_{xx} = \sigma_{zz} = \sigma_p$, $\sigma_{yy} = \sigma_o$, and $\sigma_{xz} = -\sigma_{zx} = \sigma_h$ where $\sigma_p$ is the Pedersen, $\sigma_o$ is the parallel and $\sigma_h$ is the Hall conductivity. Equations (5) and (6) are Maxwell's equations for the electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$. Using for the zeroth-order density $\rho_o$ an isothermal model, that is to say

$$\rho_o = \rho_{oo}\, e^{-z/H}, \tag{7}$$

with

$$H = \frac{p_o}{g\rho_o} = \frac{c_s^2}{\gamma g}, \tag{8}$$

where $c_s$ is the speed of sound and $p_o$ is the zeroth-order pressure, linearisation of (1) to (4) yields

$$\frac{\partial^2(\rho_o\mathbf{u}_1)}{\partial t^2} = c_s^2\nabla\left[\nabla\cdot(\rho_o\mathbf{u}_1)\right] + g(\gamma-1)(\nabla\cdot(\rho_o w_1)\mathbf{1}_z + g\left[\nabla\cdot(\rho_o\mathbf{u}_1)\right]\mathbf{1}_z +$$
$$+ \left[\left\{\frac{\vec{\sigma}}{\rho_o}\cdot\left[\rho_o\frac{\partial \mathbf{E}_1}{\partial t} + \frac{\partial(\rho_o\mathbf{u}_1}{\partial t}\times\mathbf{B}_o\right]\right\}\times\mathbf{B}_o\right] \tag{9}$$

where we assume $\mathbf{B}_o = B_o\mathbf{1}_y$ and $\mathbf{u}_o = 0$. Equations (4) to (6) yield

$$\nabla\cdot(\nabla\times\mathbf{E}_1) + c^{-2}\frac{\partial^2\mathbf{E}_1}{\partial t^2} + \mu_o\frac{\vec{\sigma}}{\rho_o}\cdot\left[\rho_o\frac{\partial \mathbf{E}_1}{\partial t} + \frac{\partial(\rho_o\mathbf{u}_1)}{\partial t}\times\mathbf{B}_o\right] = 0. \tag{10}$$

For vanishing conductivity (9) reduces to the wave-equation for acoustic gravity waves in a neutral atmosphere and (10) to the equation for electromagnetic waves. we now take $\sigma_p/\rho_o$ and $\sigma_h/\rho_o$ to be height-independent which, beyond 100 km height, seems a fare assumption according to the data compiled by Verniani (Verniani, 1974). In that case (9) does not contain height-dependent coefficients anymore. Ignoring the remaining height-dependent coefficient in the last term of (10) we perform a Fourier-Laplace transformation on $\mathbf{E}_1$ and

$\rho_o u_1$. This implies that the results will have significance only in some W.K.B. sense. Equations (9) and (10) can then be written as

$$\begin{bmatrix} A & C \\ D & B \end{bmatrix} \begin{Bmatrix} x \\ y \\ z \\ \rho_o u \\ \rho_o v \\ \rho_o w \end{Bmatrix} = 0$$

$$\{A\} = \begin{bmatrix} k_y^2+k_z^2-i\mu_o\sigma_p\omega-\omega^2/c^2 & -k_xk_y & -k_xk_z-i\mu_o\sigma_h\omega \\ -k_xk_y & k_x^2+k_z^2-i\mu_o\sigma_o\omega-\omega^2/c_1^2 & -k_yk_z \\ -k_xk_z+i\mu_o\sigma_h\omega & -k_yk_z & k_x^2+k_y^2-i\mu_o\sigma_p\omega-\omega^2/c^2 \end{bmatrix}$$

$$\{B\} = \begin{bmatrix} \omega^2-k_x^2c_s^2+i\omega\omega_p & -k_xk_yc_s^2 & -k_xk_zc_s^2+i\omega\omega_h+ig(\gamma-1)k_x \\ -k_xk_yc_s^2 & \omega^2-k_y^2c_s^2 & -k_yk_zc_s^2+ig(\gamma-1)k_y \\ -k_xk_zc_s^2-i\omega\omega_h+igk_x & -k_yk_zc_s^2+igk_y & \omega^2-c_s^2k_z^2-ig(\gamma-1)k_z+i\omega\omega_p-igk_z \end{bmatrix} \qquad (11)$$

$$\{C\} = \begin{bmatrix} -i\sigma_2\omega B_o/\rho_o & 0 & i\sigma_1\omega B_o/\rho_o \\ 0 & 0 & 0 \\ -i\sigma_1\omega B_o/\rho_o & 0 & -i\sigma_2\omega B/\rho_o \end{bmatrix}$$

$$\{D\} = \begin{bmatrix} -i\mu_o\sigma_2\omega B_o & 0 & i\mu_o\sigma_1\omega B_o \\ 0 & 0 & 0 \\ -i\mu_o\sigma_1\omega B_o & 0 & -i\mu_o\sigma_2\omega B_o \end{bmatrix}$$

where $\omega$ is the wave-frequency and $k_x$, $k_y$, and $k_z$ are the wave-numbers in x, y, and z-direction and where $\omega_p = \sigma_p B_o^2/\rho_o$ is the Pedersen damping rate and $\omega_h = \sigma_h B_o^2/\rho_o$ is the Hall frequency .

It is generally assumed that when $\sigma_p$ and $\sigma_h$ are small enough, the determinant of the matrix in (11) can be approximated by the product of the determinants A and B (Hines, 1955). This gives the following dispersion equation

$$\|A\| \cdot \|B\| = 0. \qquad (12)$$

Whether this approximation is valid or not for the real atmosphere is studied by comparing dispersion curves for acoustic gravity waves calculated from the full determinant of (11) with the ones calculated using (12).

Therefore we first evaluate $\|B\|$ which results in

$$\|B\| = \omega^2\left\{\Omega^4 - \Omega^2c_s^2(k_x^2+k_z^2+ik_z/H) + \omega_b^2c_s^2k_x^2 + c_sk_x\omega\omega_h\omega_b\frac{2-\gamma}{\sqrt{\gamma-1}} - \omega^2\omega_h^2\right\}$$
$$+ c_s^2k_y^2(\Omega^2\omega_b^2+\omega^2\omega_h^2-\Omega^4) = 0 \qquad (13)$$

where

$$\Omega^2 = \omega^2 + i\omega\omega_p, \tag{14}$$

and

$$\omega_b = g\sqrt{\gamma-1}/c_s \tag{15}$$

is the Brunt-Väisälä frequency. Because of (12), $\|B\| = 0$ can be interpreted as a dispersion relation. It can also be obtained directly from (7) when $E_1$ is taken to be zero. This implies that this dispersion relation is not based on some W.K.B. approximation, because (7) does not have any height-dependent coefficients. For acoustic gravity waves in a neutral atmosphere we have $Im(k_z) = (2H)^{-1}$ (*Hines*, 1960). Because we want to study the effects on dispersion caused by conductivity compared to the neutral atmosphere, we substitute this in the dispersion relation $\|B\| = 0$.
This yields

$$\omega^2\left\{\Omega^4 - \Omega^2 c_s^2\left[k_x^2+k_z^2+\omega_a^2/c_s^2\right] + \omega_b^2 c_s^2 k_x^2 + c_s k_x \omega\omega_h\omega_b\frac{2-\gamma}{\sqrt{\gamma-1}} - \omega^2\omega_h^2\right\} +$$

$$+ c_s^2 k_y^2(\Omega^2\omega_b^2+\omega^2\omega_h^2-\Omega^4) = 0 \tag{16}$$

where

$$\omega_a^2 = \frac{c_s^2}{4H^2}, \tag{17}$$

and $k_x$, $k_y$, and $k_z$ are real wave-numbers. Equation (16) contains odd powers of $\omega$ and $k_x$. This reflects the presence of a $\mathbf{g}\times\mathbf{B}_o$ drift with a velocity $g/\omega_h$.

To investigate (16) further, we restrict ourselves to waves for which the wave-vector $\mathbf{k}$ has only a component in the x, the y, or the z-direction.

For $\mathbf{k} = k_x\mathbf{l}_x$ (16) reduces to

$$\Omega^4 - \Omega^2 c_s^2 k_x^2 - \Omega^2\omega_a^2 + \omega_b^2 c_s^2 k_x^2 + c_s k_x \omega\omega_h\omega_b\frac{2-\gamma}{\sqrt{\gamma-1}} - \omega^2\omega_h^2 = 0 \tag{18}$$

For $\mathbf{k} = k_y\mathbf{l}_y$ we find

$$\omega^2\Omega^4 - \Omega^4 c_s^2 k_y^2 - \omega^2\Omega^2\omega_a^2 + \Omega^2 c_s^2 k_y^2\omega_b^2 - \omega^4\omega_h^2 + \omega^2 c_s^2 k_y^2\omega_h^2 = 0, \tag{19}$$

and for $\mathbf{k} = k_z\mathbf{l}_z$ we obtain

$$\Omega^4 - \Omega^2 c_s^2 k_z^2 - \Omega^2\omega_a^2 - \omega^2\omega_h^2 = 0. \tag{20}$$

When $\omega_h \ll \omega$, there is no difference between (16) and the dispersion equation for the *neutral atmosphere* except that $\Omega$ has to be replaced by $\omega$. Looking at (18) and (20) this means damping of the acoustic gravity waves in x- and z-direction with a typical damping rate $2/\omega_p$. Analysing (19) we find that for $\omega \ll \omega_b$ and $\omega \gg \omega_a$ the waves in the y-direction are undamped and that for $\omega \simeq \omega_b$ and $\omega \simeq \omega_a$ the waves are damped again with a damping rate $2/\omega_p$. It is interesting to know whether the exponential growth of the wave amplitude with height goes faster of slower than the exponential damping in time.

Therefore we compare $(2Hk_z)^{-1}$ with $\omega_p(2\omega)^{-1}$. For characteristic values of the parameters ($H = 10^4$m, $\omega = 10^{-5}$-$10^{-4}$s$^{-1}$, $\omega_p = 10^{-3}$-$10^{-2}$s$^{-1}$, $k_z = 10^{-3}$-$10^{-5}$m$^{-1}$) this means that the exponential growth with height is almost always the most important effect. Nevertheless damping is not negligible, especially for horizontally propagating waves. When exponential growth with height is the more important effect, non-linear behaviour is to be expected for upward propagating waves, (Miesen, 1988).

When the Hall frequency is not much smaller than the Brunt-Väisälä frequency, also the real part of $\omega$ is influenced (when $\omega_h \ll \omega$ only the imaginary part of $\omega$ is changed due to conduction), but also coupling with electromagnetic modes then becomes more important.

### 3. Results and conclusions

We have calculated curves representing the dispersion of acoustic gravity waves for characteristic values of the parameters at several heights in the atmosphere. We assumed the horizontal wave-vector $(k_x.k_y)$ to be real-valued and the frequency $\omega$ to be complex. The imaginary part of the vertical wave-number was assumed to be $(2H)^{-1}$as is the case for the neutral atmosphere. For realistic values of the parameters the calculations show no significant difference between the real part of the frequency whether the conduction is incorporated or not.

In figure 1 we compare curves for the imaginary part of the frequency versus $k_x$, $k_y$, and versus the real part of the vertical wave-number $k_z$. The curves resulting from the dispersion equation when hydrodynamic-electromagnetic coupling is neglected show the characteristics as discussed after (20), i.e. acoustic gravity waves perpendicular to the earth's magnetic field are damped with a typical damping rate $\omega_p/2$ and the principal wave parallel to the earth's magnetic field is only damped when its frequency is approximately equal to the Brunt-Väisälä or to the acoustic cut-off frequency.

The curves calculated using the dispersion equation for waves with hydrodynamic-electromagnetic coupling included show no significant difference for waves propagating parallel to the earth's magnetic field and for waves propagating in vertical direction. However, the curves for gravity waves propagating horizontally and perpendicular to the earth's magnetic field differ. These curves are characteristic for heights beyond 120 km, and differ from the predictions made by Hines (Hines, 1955), who found all acoustic gravity waves to be damped with a damping rate $\omega_p/2$.

We found an exception for a height of 100 km where the quotient of the Brunt-Väisälä and the Hall frequencies has its maximum. At this height the real and the imaginary part of the frequency differ significantly when calculated from the two dispersion equations. When hydrodynamic-electromagnetic coupling is included the cut-off frequency for the acoustic branch is smaller. Consequently the frequency band where propagation is not possible disappears. We will come back to this and to the difference in damping rate for gravity waves propagating perpendicular to the earth's magnetic field incorporating coupling or not in a forthcoming publication.

Further investigations showed that the quotient of the Hall frequency and the Brunt-Väisälä frequency is indeed an important parameter that determines the influence of the hydrodynamic-electromagnetic coupling on dispersion of acoustic gravity waves. In conclusion we can say that gravity waves are damped at a rate equal to $\omega_p/2$ when they propagate parallel to the earth's magnetic field with a frequency near to the Brunt-Väisälä frequency. Gravity waves with frequencies much smaller than the Brunt-Väisälä frequency are not damped. The damping of gravity waves propagating perpendicular to the earth's magnetic field needs further investigation because hydrodynamic-electromagnetic coupling seems to play an important role. Acoustic waves are damped in the ionosphere at a rate equal to $\omega_p/2$, except when they propagate parallel to

the earth's magnetic field and with a frequency larger than the acoustic cut-off frequency. This is in contradiction with Hines's results (Hines, 1955).
Though the waves are damped due to conductivity, their amplitude grows with height- due to the exponential decrease of the mass density- as long as the vertical wave number is smaller than $\omega/(\omega_p H)$. These conclusions are correct as long as $\omega_h/\omega_b < 10^{-2}$ which is true except for a height of approximately 100 km at daytime (Cole and Pierce, 1965).

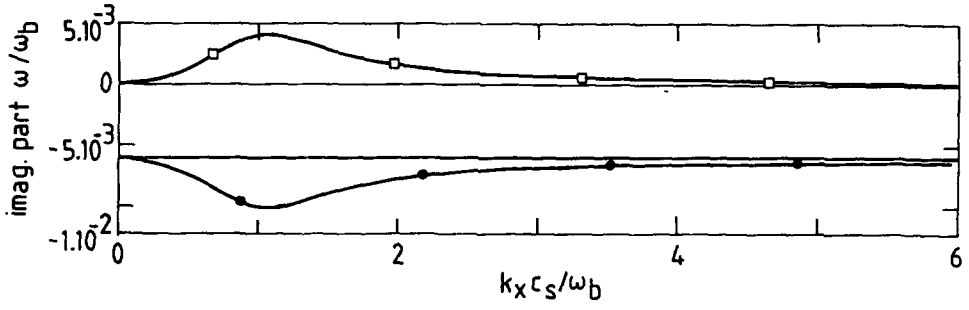



*FIGURE 1a: Imaginary part of $\omega/\omega_b$ versus $k_x$.*

*The horizontal line at $-\omega_p/(2\omega_b)$, is the damping rate for acoustic gravity waves propagating in the x-direction when hydrodynamic-electromagnetic coupling is neglected. The curve with the dots gives the damping rate for acoustic waves and the curve with the squares gives the damping rate for gravity waves both when hydrodynamic-electromagnetic coupling is included. These curves have been calculated for values of the relevant parameters that are characteristic for a height of approximately 200 km (Banks and Kockarts, 1973 and Verniani, 1974): $\gamma = 1.4$, $g = 9.5 m/s^{-2}$, $H = 5 \cdot 10^4 m$, $C_s = 815 m/s$, $\omega_b = 7.4 \cdot 10^{-3} s^{-1}$, $B_0 = 3.2 \cdot 10^{-5} T$, $\sigma_0 = 10^{1.5} mhos/m$, $\omega_p = 9 \cdot 10^{-5} s^{-1}$, $\omega_h = 3 \cdot 10^{-6} s^{-1}$.*

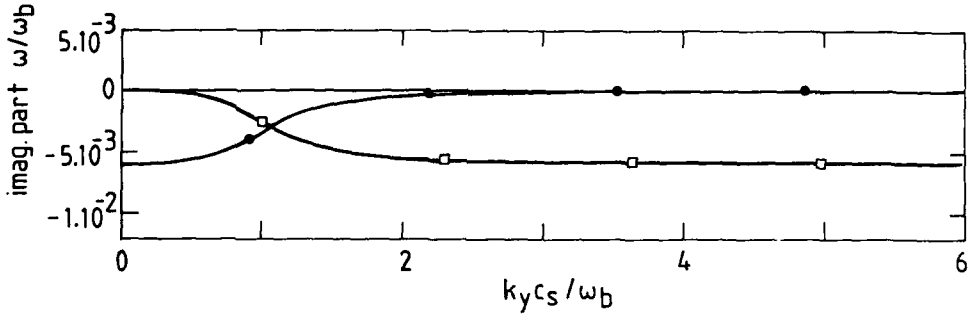



*FIGURE 1b: Imaginary part of $\omega/\omega_b$ versus $k_y$.*

*These curves give the damping rate for the acoustic waves (dots) and gravity waves (squares) that propagate in the y-direction with and without hydrodynamic-electromagnetic coupling. For this direction of propagation there turns out to be no significant influence of this coupling; the curves coincide.*

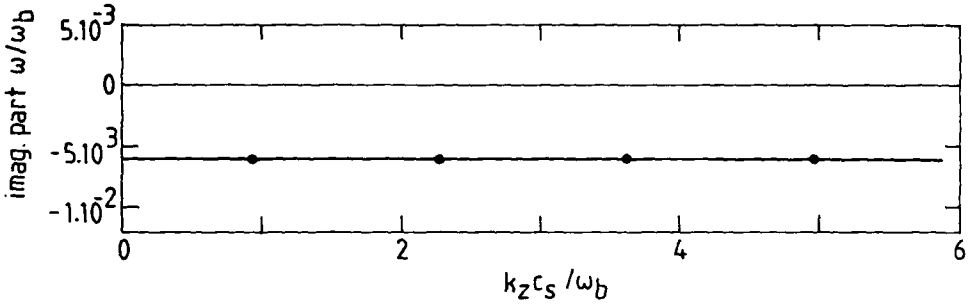



FIGURE 1c: *Imaginary part of* $\omega/\omega_b$ *versus* $k_z$.

*Since gravity waves can not propagate purely vertically, only the damping rate for the acoustic waves has been plotted in this figure. Here also it turns out to make no significant difference whether coupling is included or not.*

**References**

Banks, P.M. and Kockarls, G., Aeronomy, Part A and B, Academic Press, New York, 1973.

Cole, Jr., R.K. and Pierce, E.T., Electrification in the earth's atmosphere for altitudes between 0 and 100 kilometers, J. Geoph. Res., 70, 2735, 1965.

Hanson, W.B., Structure of the ionosphere. In: Satellite Environment Handbook (F.S. Johnson, Editor), Stanford University Press, Stanford, Calif., 187 pp, 1965.

Hines, C.O., Generalized magneto-hydrodynamic formulae, Proc. Cambridge Phil. Soc., 49, 299, 1953.

Hines, C.O., Hydromagnetic resonance in ionospheric waves, J. Atm. Terr. Phys, 7, 14, 1955.

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

Hines, C.O., The upper atmosphere in motion, American Geophysical Union, Washington, 115-469, 1974.

Miesen, R.H.M., Non-Linear wave-equations for low-frequency acoustic gravity waves, J. Atmos. Scien., to be published, 1988.

Verniani, F., Structure and dynamics of the upper atmosphere, Elsevier, Amsterdam, 119, 1974.

## DISCUSSION

**K.C.Yeh**

I would like to first congratulate you on working on this difficult coupling problem between electromagnetic waves and gravity waves. You have shown that as the result of this coupling what effects could be expected on gravity waves. Have you also looked at the result of this coupling on the propagation of electromagnetic waves?

**Author's Reply**

Thank you for your congratulations. As to your question, the effects of the coupling for the electromagnetic waves are also a very interesting topic which, however, we didn't investigate yet. We hope to do so in the future.

# ATMOSPHERIC GRAVITY WAVES IN THE AURORAL SOURCE REGION OF TIDs, STUDIED WITH THE EISCAT CP2-DATA

**Mauelshagen, H.-P. and Schlegel, K.**
Max-Planck-Institut für Aeronomie, D3411 Katlenburg-Lindau, FRG

## ABSTRACT

EISCAT offers the possibility to measure simultaneously the electron density, electron temperature, ion temperature and ion drift velocity in the height range from about 80 km to around 600 km. In the CP2-mode the EISCAT radar looks in four different directions within every 6 minutes. Thereby it receives data along the four edges of an oblique topstanding pyramid near Tromsø (Norway), that is inclined a little bit to the south-east. In each direction there are data from 21 height gates between approximately 130–560 km and additionally there are data from Kiruna (Sweden) and Sodankylä (Finland) from one height at about 278 km which allow the estimation of the three-dimensional ion drift vector in the four points of intersection at this heigth. The EISCAT facility was operated in this mode since 1984 for over 600 hours which produced time series of every measured ionospheric parameter and in any of the 84 locally spaced scattering centers.

This paper is a preliminary report of the study of atmospheric gravity waves and TIDs with these data where we just present the methods of estimating different TID-events, their direction and our further aims of study. For this purpose we confine ourselves more or less to a magnetically very quiet 38-hour period on the 5.–6. May 1987. Contrary to more active days it is possible in this period to show not only different TID-events but also to resolve a frequency-time evolution of sudden events. For this very weak magnetic activity it is unlikely that the present TIDs originate in auroral effects. Thereby one can work out criteria to distinguish events of propagating waves and of waves which are just beginning. Summing up this work shows that it is possible to estimate the whole complex wavenumber as a function of time of every resolvable wave-like structure. This will now offer a possibility to test the theories of the generation, the propagation and the dissipation of TIDs.

## INTRODUCTION

Travelling ionospheric disturbances (TIDs) are nowadays known to be the response of the ionized part of the ionosphere on gravity waves in the neutral background gas of the upper atmosphere. But nonetheless there remain many open questions concerning the TIDs, especially about their sources and about the kind of response of the different plasma parameters on the neutral wave. Historically all TIDs have been thought to come from the auroral zones and move more or less in the direction of the equator. Today its known that only the large-scale ones (periods of 30 min – 3 h and speeds of 400–1000 m/sec) behave this way while the medium-scale ones (periods of 15–60 min and speeds of 100–250m/sec) are now seen to come up from the lower and middle atmosphere with meteorological sources in the troposphere and with directions commonly opposite to the neutral wind direction.

Soon after the famous paper of *Hines* (1960)[1] there began a very active phase of study in this part of ionospheric physics as seen for example in the review of *Francis* (1975)[2] and the enormous number of papers cited in it. Most of this work can be clearly divided into either observational or theoretical studies with hardly any overlap of these two. Despite, the quality of these old papers is very high and is rarely reached by newer ones (see for example *Bowman* (1968)[3], *Georges* (1968)[4], *Klostermeyer* (1969, 1972)[5,6], *Hooke* (1968)[7] or *Volland* (1969*a*, *b*)[8,9]). Also some books appeared (A collection of *Hines*'s work (1974)[10], *Beer* (1974)[11], *Gossard and Hooke* (1975)[12] and *Yeh and Liu* (1972)[13]), describing the thereby reached state of knowledge in a reviewlike manner. After that the general interest in this theme faded for about 10 years although some few important papers appeared (for example *Richmond*, 1978[14]).

In the last few years there is again a growing engagement in this field, but in comparison with the older work the methods have totally changed. It were mainly these new methods that made it necessary to repeat similar observational work that had been done 20 years ago. On the other hand the open questions today are nearly the same as the remaining problems of the older days (see for example *Hunsucker*, 1982[15] and *Williams et al.*, 1988[16]). But now it should be possible to use the enormous data sets and the developed computer resources to solve these problems and to obtain both better observational and out of this better theoretical insight in the matter of gravity waves in the upper atmosphere. A special and very promising facility for such studies will be the EISCAT radar. This paper is a preliminary report of such work and describes the beautiful new possibilities one has with EISCAT.

## THE CP2–MODE OF EISCAT

EISCAT the European incoherent scatter facility is a tristatic radar situated in northern Scandinavia. The EISCAT-UHF system used here consists of three identical parabolic antennas from which one is used for transmitting and receiving while the other to are just receiving antennas. The transmitting antenna is based near Tromsø (Norway) and the other two are located in Kiruna (Sweden) and in Sodankylä (Finland). In contrast to other incoherent scatter radars these three EISCAT antennas offer the possibility to study the ionospheric parameters from all three antennas simultaneously

at the same place and the same time. This gives the full three dimensional ion drift velocity vector and allows studies of temperature anisotropies. EISCAT is now operated in different modes. Besides three common modes there is operation time for each contributing country to do some own special measurements. The CP2–data used here result from one of the common programs which is run usually once a month. EISCAT was operated in this mode for over 600 hours since 1984 recording mostly for time periods of 24 to 38 hours.

Measuring along one direction and moving to the next direction lasts about 1 to 2 minutes. Sometimes, however it is possible to resolve some ionospheric data from integration times of just 10 sec. In order to get some different directions and an acceptable time resolution too, four measuring directions have been chosen for the CP2-mode. The four directions are vertically upwards, to the south, to the south-east and magnetically upwards. The cycle through these four positions is repeated every 6 min. In this mode the two remote sites of EISCAT are just looking at one height of every beam. This height of the common volume is changed in different versions of the CP2-program according to the sunspot cycle and is about 278 km in the recent version.

Beside the data from the four common volumes at this one height EISCAT offers the data recorded along the Tromsø antenna beam. This means that in the CP2–mode the Tromsø antenna is looking along the four edges of an topstanding oblique pyramid inclined a little bit to the south-east. The normal EISCAT CP2–data from Tromsø consist of the ionospheric parameters of 21 height gates from each of the four measuring directions which cover a height range of about 130 to 560 km. Additionally there exist data from about 60 to 150 km with a height resolution of just 3 km which are obtained with a multipulse technique (see for example *Lehtinnen and Huuskonen* 1984[17]). In this preliminary part of the study we have just used the long–pulse EISCAT–data because there are great data lags in the multipulse data and because we are interested mainly in the occurence of TIDs in the F–region.

Because of the different elevation angles the 21 height gates collect the data from different height intervalls in the four directions. Vertically upwards the ionospheric data are estimated every 22.5 km from 150 to 600 km while in the south–eastern direction the data are from heights of 131 to 528 km with height increments of about 20 km. Considering this and the growing horizontal distance with increasing height, one must be careful in the comparison of data from different directions. We therefore have just used the data from the common volumes at the height of about 278 km for interdirectional studies and even for this values the data in the single directions have to be interpolated. Despite these systematic geometrical difficulties and the time delay between successive measurements one has time series of four ionospheric parameters in each of the 84 locally spaced scattering volumes which principally can be correlated to each other. Moreover the locally different electric field from the four measured threedimensional ion drift velocity vectors can be estimated. With this and the time resolution of 6 min in every time series in mind it is clear that these CP2–data from EISCAT offer a wonderful possibility to study atmospheric gravity waves and TIDs.

## THE IONOSPHERIC PARAMETERS ON THE 5.–6. MAY 1987

As an example of the general behaviour of the EISCAT-data we present here the height profiles of the estimated ionospheric parameters on 5.–6. May 1987. For every parameter the values for one of the four directions is shown, the electron density, electron temperature and ion temperature are explained by their vertical profiles, while for the ion drift velocity the direction parallel to the magnetic field is used to suppress the effect of electric fields. In order to give clearer height profiles the estimated 16 data points along the used height range are not connected linearly but with the help of a polynomial interpolation. In the time direction the data were shown in the form of running averages of 5 points, which means that at any point the data were averaged over a time period of ±15 min. These data from the 5.–6. May 1987 represent a magnetic quiet period with $\sum$Kp being only 9- and 9+ respectively. Therefore Figures 1-4 characterize the averaged quiet day behaviour of the auroral upper ionosphere.

At Tromsø the local time is $UT + 1.5\ h$ and the magnetic local time is $UT + 2.5\ h$. Therefore the local midnight and local noon are at $22^{30}UT$ and $10^{30}UT$ while the magnetic local midnight and noon are at $21^{30}UT$ and $9^{30}UT$ respectively. On the ground, the sunset and sunrise at Tromsø in the night from 5.–6. May 1987 were at about $20^{30}UT$ and $1^{00}UT$. With increasing height the time delay between these points decreases rapidly until a height of around only 30 km it vanishes totally. This means that above 30 km the atmosphere is sunlit throughout the night. But it has to be kept in mind that the angle of incidence of sunlight is going up in the evening and that at all times between the sunset and the sunrise at the ground the incoming sunlight at the ionosphere is shining from beneath that layer what means that the ionizing radiation is filtered out previously.

In the daytime the height of maximum electron density is as low as 270 km and at this height range electron density seems to go up and down opposite to the zenith angle. Above 350 km on the other hand there are no daytime changes visible. The estimated electric field data show that the enhancements of the electron density at lower heights at about 20 UT on the 5. and about 18–20 UT on the 6. May are related to higher electric fields. While at other times the electric field is lower then about 2 mV/m it goes up to 20 mV/m between 19 and 20 UT on the 5. May and rises at times to over 30 mV/m after about 16 UT on the 6. May. Despite these electric fields are not very large they affect the situation in the otherwise very quiet conditions.

An effect of the sometimes enhanced electric field can be recognized in the ion temperature too, in form of an increased temperature throughout all heights as shown in Figure 2. In general the ion temperature seems to rise with height. The outstanding features of this figure are therefore the maxima of ion temperature at both days in the height range of 200 to 220 km in the daytime. At greater heights the ion temperature rises again above about 300 km. Beside the existence of these daytime maxima there are other enormous day night differences. Especially one should mention the different day night behaviour in the lower and upper heights. Below about 250 km the ion temperature goes up and down with incoming

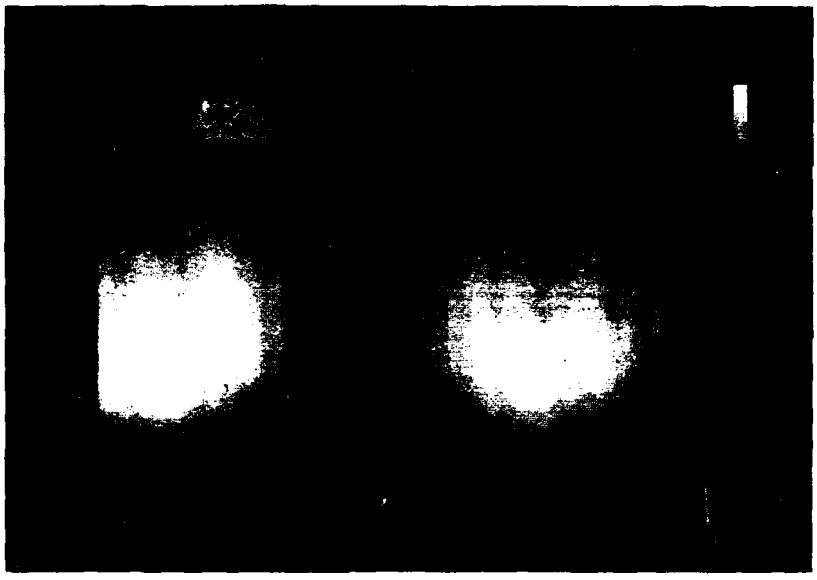


*Figure* 1: Vertically measured electron density during the 38 hour magnetically quiet time period of the 5.–6. May 1987. The data are smoothed in time of the order of half an hour and the units are $10^{11}$ electrons per cubic meter.

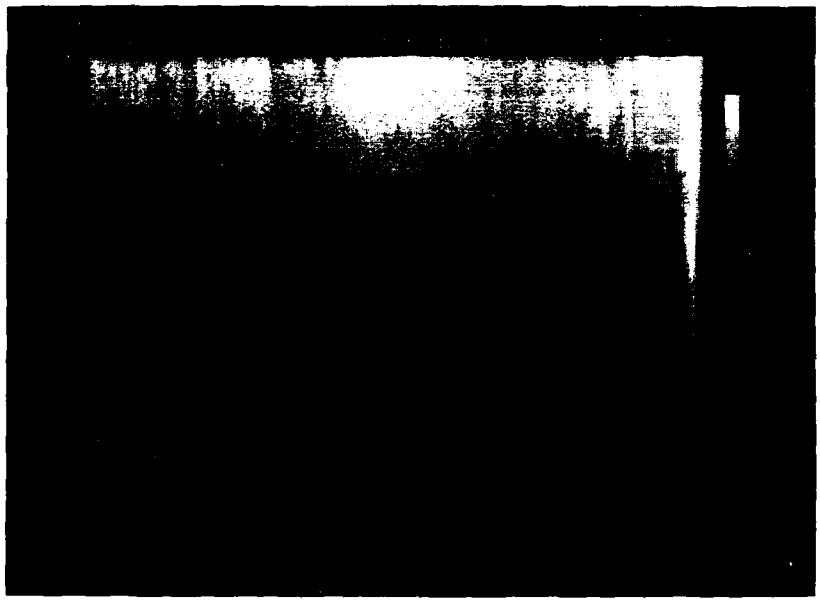


*Figure* 2: Vertically measured ion temperature during the 38 hour magnetically quiet time period of the 5.–6. May 1987. The data are smoothed in time of the order of half an hour and are given in degrees Kelvin.

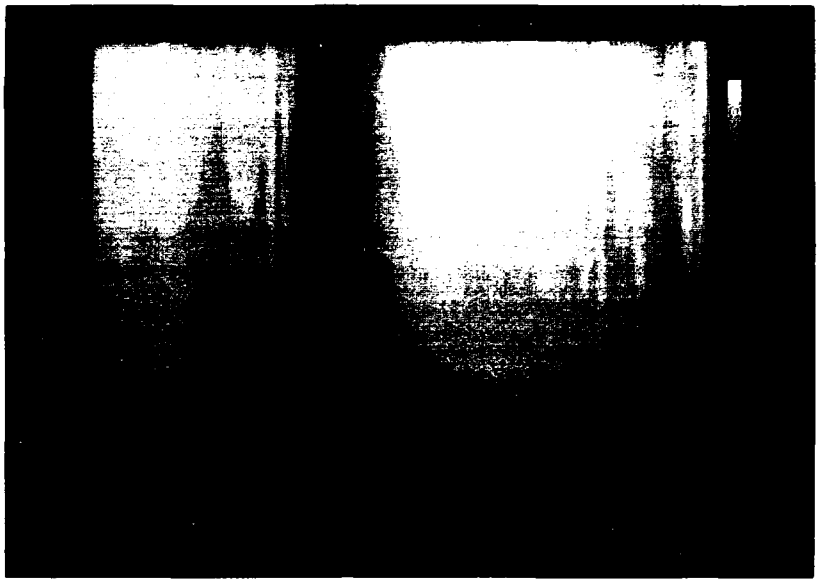


*Figure* 3: Vertically measured electron temperature during the 38 hour magnetically quiet time period of the 5.–6. May 1987. The data are smoothed in time of the order of half an hour and are given in degrees Kelvin.



*Figure* 4: Ion drift velocity parallel to the magnetic field measured during the magnetically quiet time period of the 5.–6. May 1987. The data are smoothed in time of the order of one hour and are given in meters per second with positive velocities going upwards and negative velocities downwards.

solar irradiation as it is expected. Above about 300 km however the variation has the opposite sign. This can be explained perhaps by an expanding and a compressing of the neutral thermosphere in the daily course of the solar heating.

The behaviour of the electron temperature (Figure 3) is totally different. It seems as if the electron temperature is directly affected by the solar ionizing radiation as long as electric fields are sufficiently low. An argument for this is the fact that the period of reduced electron temperature during the night of 5. to 6. May decreases with increasing height. During the next night this can not be seen because the particle precipitation associated with the increased electric field causes a heating of the electrons. During daytime the electron temperature in the F-region is about twice the ion temperature. On the other hand the electron temperature shows no maximum with height but is increasing more or less continously.

Figure 4 shows the gray-scale coded height profile of the ion drift velocity parallel to the magnetic field. These data were smoothed by a running average over 10 data points which means a temporal moving average of one hour. In the height direction such an average is insignificant because the few existing data points are only arbitrarily enhanced by polynominal interpolation. Despite this averaging these ion drift velocity data are more variable then data of the other ionospheric parameters. This is due to the different effects that can influence this drift velocity. Among others there are up and downward neutral winds, electric fields and horizontal meridional neutral winds that push the ions upwards or downwards along the magnetic field. Beside the times of the slightly enhanced electric field it seems as if the ion drift is upward in the night and downward through the day in nearly all heights. At the greatest heights however it is nearly continously upwards and in the lowest heights it is more or less undefined. The upper drifts may be explained by the global mean meridional circulation which should be upwards and southwards in the summer northern thermosphere (see for example *Richmond* 1983[18]). The ion drifts in the other heights might then be explained by a presumed southward neutral wind at night and a northward neutral wind during the daytime.

## EXAMPLES OF WAVE-LIKE STRUCTURES

To get familiar with TIDs and atmospheric gravity waves let us first show a few examples of wave-like structures. In the gray-scale coded smoothed height profiles of the ionospheric parameters shown in Figures 1–4 there are just some wave-like structures to be seen. For example in Figure 1 some isolines of the electron density behave in a wave-like manner with an amplitude of over 20 km and periods of mainly 2–6 hours. It is now necessary to decide which of these structures are really due to atmospheric gravity waves and which are just an effect of the diurnal change of solar UV-incidence or

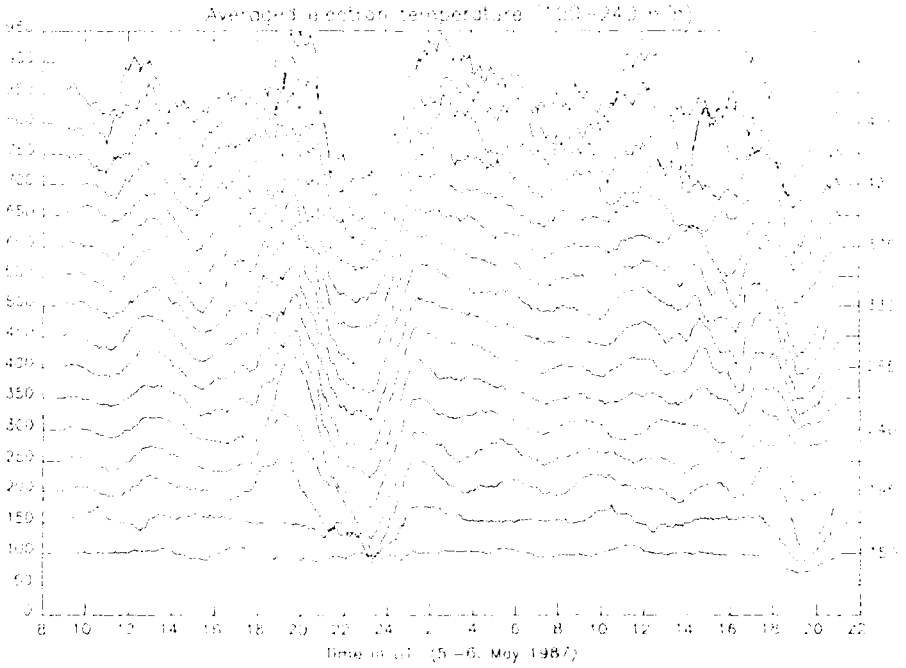



*Figure 5:* Filtered data from 2 hours to 4 hours of the 16 lowest vertically measured electron temperatures over Tromsø. Each line consists of the values from one height and the data from successive heights are shifted

which are connected to electrodynamic phenomena. In order to distinguish such similar looking but different wave-like events it is first of all useful to estimate the electric field, but then it is especially helpful to study their height dependence. Finally one can compare their appearence in the different parameters as in the different directions by plotting the data with a maximum resolution of the events.

Beside gray-scale coded pictures, or contoures one can present the height and time dependent data in the form of intensity lines from different heights with intensity shifts between them representing the height differences. These quasi-three dimensional pictures are sometimes more impressive and more illustrative than gray-scale coded pictures. Figure 5 is an example of such a plot which shows the 2–4 hour variations in the height dependent vertically measured electron temperature. For this picture the data are first subtracted by their 4 hour moving average and are then averaged by a 2 hour moving average. The data filtered in this way are then linearly connected, though it is sometimes more useful to apply interpolation methods to give an even more impressive imagination of the considered wave-like events (see for example *Schlegel* 1986[19]).

The appearing difficulties in deciding whether an event is caused by a neutral wave or is just an effect of other natural phenomena in the polar ionosphere gets clear also in Figure 5. The outstanding decrease in electron temperature through the night of the 5. to the 6. May of Figure 3 is now looking not very distinct to the wave-like event appearing a few hours earlier. Therefore one has to consider always the different kinds of data and especially the electric field to make sure that the fluctuations of the ionospheric parameters are really due to neutral waves, despite they last only for 1 to 2 waveperiods sometimes. Figure 6 shows that it is not always easy to recognize at once a wave-like event though one exspects it. The big wave-like event with a period of a little less than 4 hours in the beginning of this data set had for example been seen in the gray-scale coded tables and in the electron temperature of Figure 5. But now in Figure 6 which shows the four vertically measured ionospheric parameters at 240 km one has to look carefully to distinguish this wave from the following fluctuations in nearly the same period regime. In recognizing a neutral wave one must then consider the measuring height of the used ionospheric parameters before starting to explain sudden fluctuations in peak direction, amplitude and phase. Exact estimations of the height dependent response of the ionospheric plasma to a neutral passing wave will not be given in this preliminary study but will be studied more extensively. Figure 6 is just an example and just one case of the different behaviour of the parameters at just one height.

Another possibility is the comparison of the data from different measuring directions. In Figure 7 this is shown with the help of an other very illustrative method of presenting data with wave-like fluctuations. For this figure the electron density data at the height of 278 km are filtered with moving averages of 30 min and 6 hours. The results of these averaging processes are then plotted one above the other for each measuring direction. At the ends of the time period the data cannot be averaged within the first and last half of the averaging period. This explains the different behaviour of the plotted lines in these parts of the figure.

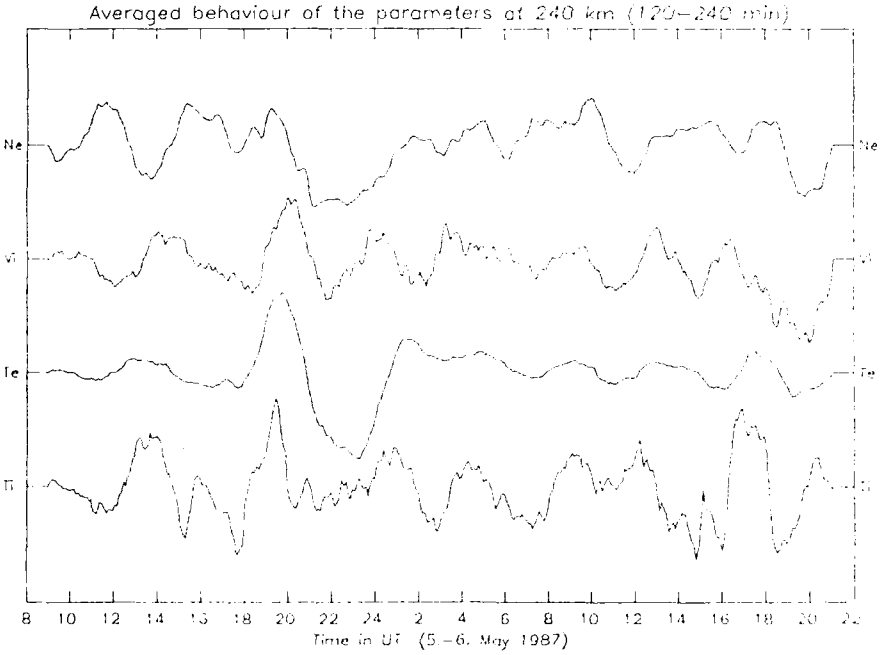



*Figure* 6: Filtered data from 2 hours to 4 hours of the vertically measured ionospheric parameters at the height of 240 km over Tromsø.

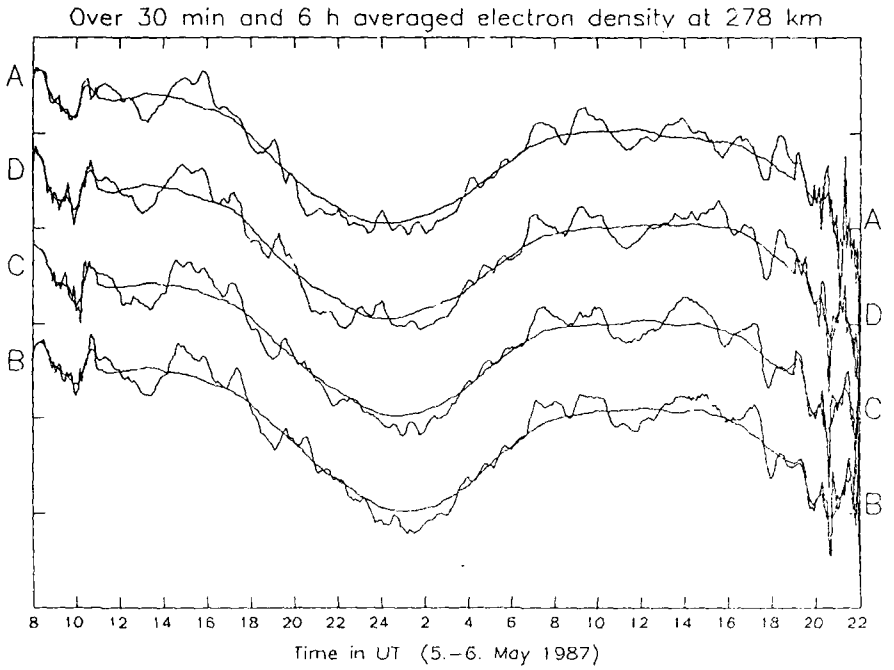



*Figure 7:* Time averages of 30 minutes and 6 hours of the electron density at about 278 km measured in the four different directions.

## RESOLUTION OF DIFFERENT WAVE-LIKE EVENTS

In order to do any further study with these data besides just describing some outstanding events one needs some special statistical methods. The problem is that despite the time series of the 84 locally spaced measuring volumes principally can be intercorrelated, the considered events last only for a few wavelengths. After a short pause there may then be an other event with nearly the same period and perhaps also with a similar propagation direction. On the other hand one wants to detect wave events in the whole period range of 20 min to 4 hours in one automatically performable program. This means

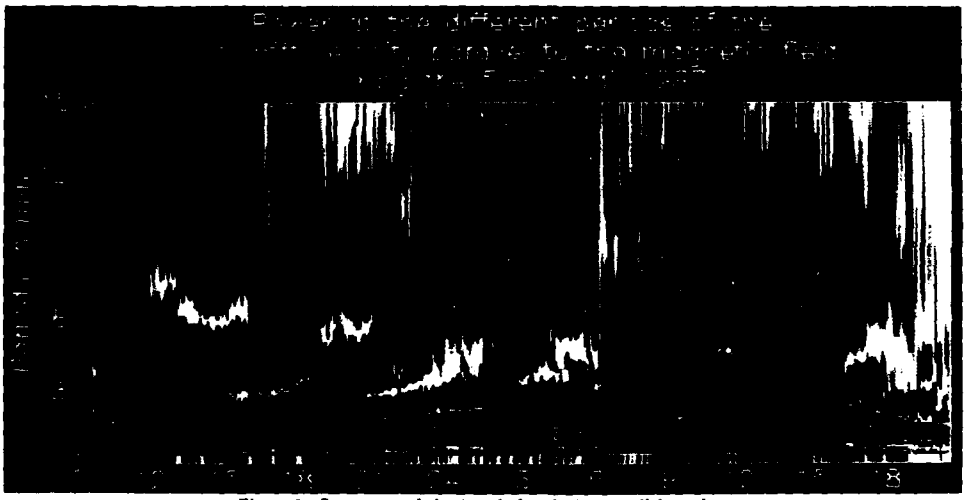


*Figure 8:* Sonogram of the ion drift velocity parallel to the magnetic field during the 5.-6. May 1987.

that in only one period of the 4 hour event there is time for 12 wavelengths of the 20 min event which then principally can be the superposition of different events in that period range. Therefore one needs a method which is able to resolve a wave that is just existing for 1 to 2 wavelengths and on the other hand one must chose only minimal time intervalls in order to prevent a mixing of different wave events.

Because Fourier analysis is insufficient to achieve such an estimation one uses here the maximum entropy method, that is superior to Fourier analysis especially for such short data sets. For an explanation of the method see for example *Ulrich and Bishop* (1975[20]) or the therein cited literature. Here we use this method in a modified form given by *Barrodale and Erickson* (1980[21,22]) which takes care especially to rounding errors and to the spectral estimation of very short data sets in respect to the estimated period. With this it is possible to arrive at a good power estimation at all periods up to four hours by taking separately three hour data sets that are obtained by shifting a 3 hour window through the whole data set by steps of one data point i.e. by 6 min. The power spectra from each of the shifted 3 hour periods are then put together in form of a sonogram (Figure 8) showing the power at any period below 150 min at all central times of the 3 hour data sets. We confined ourselves to periods less than 150 min in order to give a better period resolution in the interesting region of 30–60 min and to avoid a reduction of the power at the long period end of the spectra. On the short period end the peaks at 6 and 12 min are meaningless due to the time delay of 6 min of consecutive measuring points.

So far we have performed such sonogram construction only for the ion drift data of the direction parallel to the magnetic field in the height of 278 km but the result (for example Figure 8) is impressive. One can clearly distinguish different events and moreover some of the events show clear pattern of time evolution. This is, on the other hand, just a fortunate case that can be found only during magnetically very quiet times. Figure 9 for example, shows a similar sonogram of a magnetically just a little more active period. The $\sum$Kp are only 14+ and 25 on that days, respectively but the average power level is much higher and is of the order of the peak power level in Figure 8. This explains that it is nearly impossible to extract similar events from Figure 8, especially from these sonograms constructed from the power spectra of the ion drift.

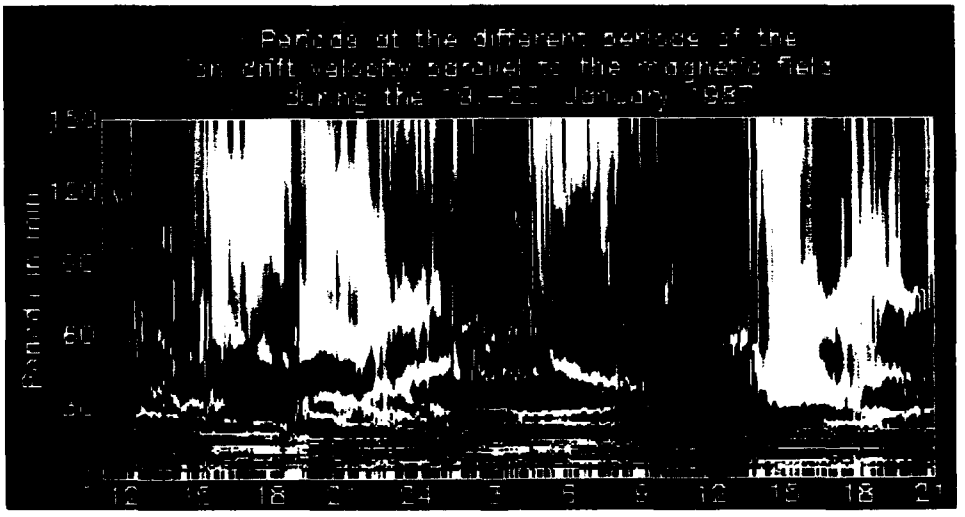



*Figure* 9: Sonogram of the ion drift velocity parallel to the magnetic field during the 19.–20. January 1987.

As a result of the sonograms one knows what periods of wave-like events are present in the data at a special time intervall. With this knowledge one can do special investigations of that time intervall in order to separate the influence of each event by applying different band filters. Figures 10 and 11 for example, show the ion drift data parallel to the magnetic field in the time interval from 18 to 24 UT on the 5. May 1987 with two different band filters. In these cases the data are first subtracted by their longer period average i.e. by their average of 36 min and 60 min respectively, and then are just smoothed by their lower band period i.e. the moving averages of 24 min and 42 min are estimated respectively. Although one can see many differences in these figures it has been found that there are additional complications due to the fact that the two main periods at about 30 min and about 60 min are nearly harmonic. In both cases, however, the tilted lines of constant phases are visible, indicating an upward moving gravity wave.

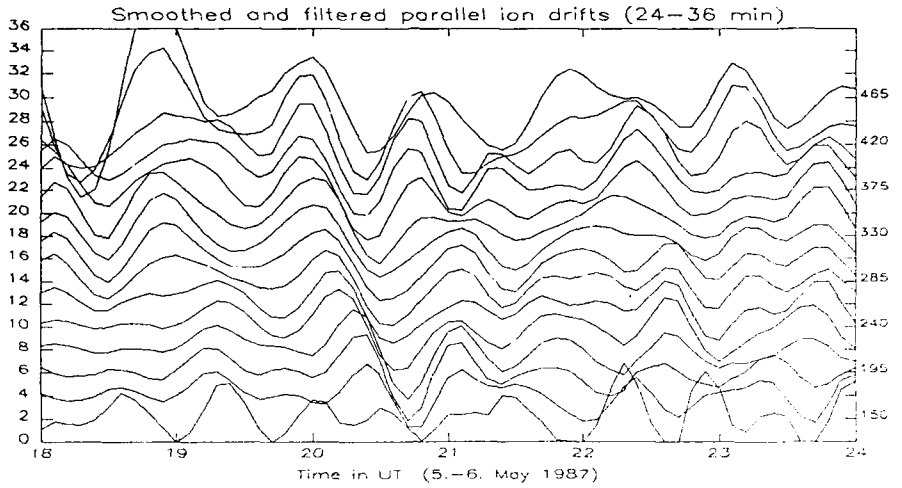



*Figure* 10: Smoothed differences of the 24 and 36 minutes averages of the ion drift velocity parallel to the magnetic field shown for 18 to 24 UT at the 5. May 1987 . Similar to Figure 5 here the data from the 15 lowest heights are plotted and shifted 2 m/sec from each other.

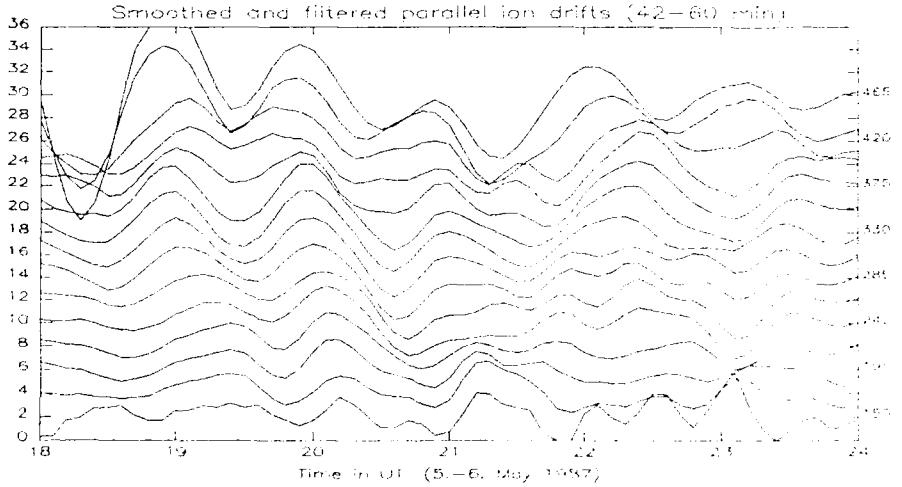



*Figure* 11: As Figure 10 but for the differences of the 42 and 60 minutes averages.

## ESTIMATION OF THE TID–DIRECTIONS

The CP2 EISCAT mode is useful to estimate the direction of atmospheric gravity waves especially because in the case of travelling waves the same wave–like structures are found in the four directions. Up to now we have used these possibilities only to estimate the direction of some TIDs out of the drift velocity of the ions. We know that considering electric field effects this is not the optimal ionospheric parameter but it is enough to show the method and some preliminary results of TID directions. The difference of this with respect to the use of other parameters should only be clearer peaks in the

correlation curves which nevertheless would be located at the same times. For clarity it should also be mentioned here that for all kinds of correlation studies it is necessary to fill eventually existing data gaps. Here we filled our data gaps by the same autoregressive methods of *Barrodale and Erickson* (1980[21,22]) which we also applied for our power estimation. This only means that we first filled all eventually existing data gaps with the same subprograms that we used afterwards to estimate the power spectra for the sonograms.

In order to estimate the more or less horizontal propagation direction we confine ourselves to just one height at the common volume at around 278 km. Fig. 12 shows an example of the constructed correlation curves of all six possible correlations of the four directions at this height. For every curve the correlation coefficient of the data set in the first direction centered at the time $t_0$ and the data set of the second direction centered at the time $t_0 + a$ were calculated for 'a' being positive and negative multiples of 6 min, and $t_0$ in the shown example is 22 UT. The resulting values of the correlation coefficients are then connected via a polynominal interpolation of 5 th order to enable artificially a better time resolution than the 6 min time delay of successive data points. Because the searched wave-like events have only a few wavelenghts duration the used data sets have only the length of about two wave periods to allow the distinction of different events and to allow changes in the propagation direction of longer events.

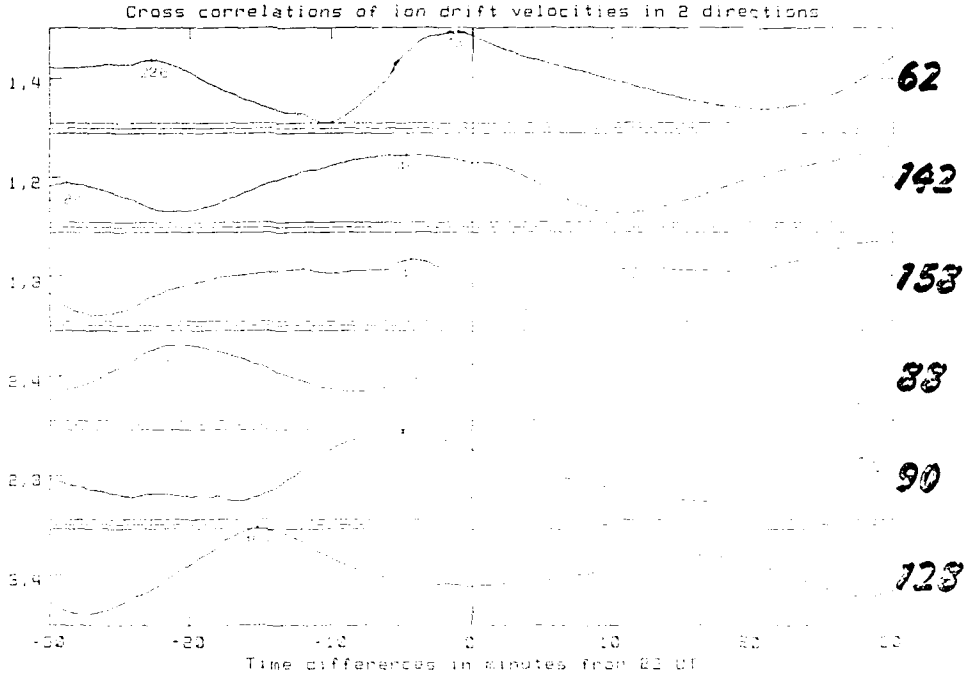



*Figure* 12: Example of the six possible correlation curves of the ion drift velocity data from the four directions at the height of about 278 km. The curves show the temporal evolution of the correlation coefficients.

The enlightening paper of *Hines* (1974)[23] gives clarity about the different speeds and velocities of waves which can be calculated. From the described correlation studies we obtain the phase trace speeds in the different directions of the measured scattering volumes. These phase trace speeds are just the quotients of the distances of the two used volumes and the time differences of the correlation maxima. For these times of maximum it actually has to be kept in mind that EISCAT measures the successive directions with time differences of 1.5 min what means that there are additional time delays of 1.5 min, 3 min or 4.5 min to be considered. Figure 12 shows that it is not always easy to estimate the time of the maximum in the correlation curves that belong to the chosen event. On the other hand one has sometimes clearly two corresponding times whose difference clearly gives the period of the wave. With the help of those six correlation curves for every event or every chosen time one usually gets 4-5 good estimates of the corresponding phase trace speeds. Because minimally two trace speeds are enough to estimate the horizontal phase trace speed and the propagation direction, this method is fully sufficient.

Though the propagation direction of TIDs can be calculated by these maximally six estimated directional phase trace speeds we have up to now just estimated the directions and speeds graphically. For this purpose we drew a sketch in the dimension of phase speeds for every event to be examined. In accordance with *Hines* (1974)[23], we estimate the direction of horizontal phase propagation by evaluation of the direction of the lowest phase trace speed of the presumed plain wave. In a sketch of the four scattering volumes we first drew one phase trace speed in arbitrary but from then on constant units.

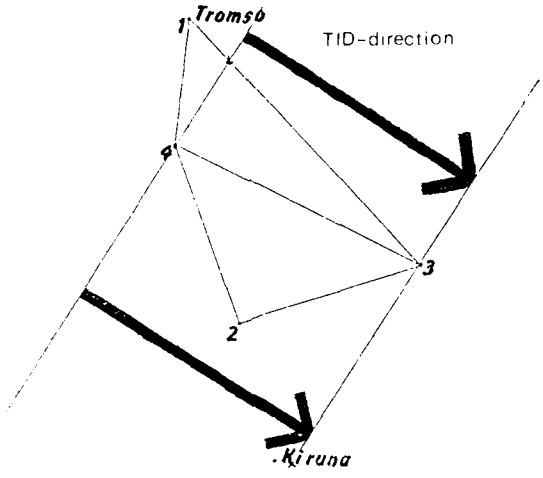



*Figure* 13: Sketch to estimate the horizontal phase trace speed and propagation direction for the same event whose correlation curves are shown in Figure 12.

We drew this speed from one of the two regarded volumes in the direction of the other under consideration of the sign. Because the units are optional it is found to be easiest to take the units with the aim that this phase trace speed just connect the two volumes. Using the adopted units of phase speed we drew now the other two phase trace speeds from each of the connected points respectively in the directions of the other two volumes taking the corresponding signs into account. From these five used phase trace speeds one can then construct a pair of parallel straight lines through the two first connected volumes and the endpoints of the sketched speeds. One can proof this construction with the sixth estimated phase trace speed which should be the distance of the two parallel lines in the direction of the two not connected volumes. The perpendicular distance of the two parallel lines is then the horizontal phase trace speed in the arbitrarily chosen units and the direction of the vertical is the horizontal propagation direction of the wave of this event.

It has to be kept in mind that these constructed parallel lines are not lines of constant phase with just one wavelength in between. As an artifical easy example let us take only three phase trace speeds that all contain the second scattering volume. First we connect the first and second volume and estimate from their range and from that trace speed the used units (m sec per cm). This choice of units fix exactly that time a wave needs to travel from volume two to volume one. At the same time period the phase of the wave travels in the directions of volume three and four exactly to that point

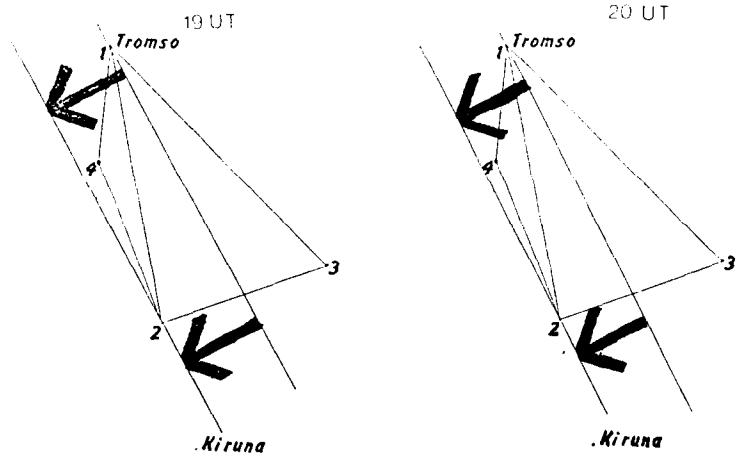



*Figure* 14: Propagation directions and phase trace speeds for the wave-like structure shown in Figure 11. The left picture shows the situation at about 19 UT and the right one an hour later.

which is given by that time and the corresponding phase trace speed. Presumed that all three speeds are positiv the three endpoints will lay on one straight line which can be seen as a line of constant phase. Normally this phase will be different from the phase of a parallel line going through volume two, but the vertical to them represent the propagation direction. A phase trace speed perpendicular to these lines is on the other hand the lowest possible existing phase trace speed which means that this is the horizontal phase trace speed in the units chosen by the first connection.

To do this construction for the correlation curves shown in Figure 12 yields Figure 13. There one sees that the propagation direction is to the south–east with a horizontal phase trace speed of 124 m/sec. Considering the wave period of 31 min one calculates the horizontal wavelength to be 231 km for this near event occuring near midnight. With this method of graphical estimation one can even study the temporal propagation behaviour of a TID. In Figure 14 for example are given the propagation directions of the filtered TID event presented in Figure 11 at two times separated by one hour. The picture shows that the situation has not changed very much within that time. The periods are 55 min and 54 min what is within the estimation error but the horizontal phase trace speeds have changed from about 40 m/sec to only around 35 m/sec. These speeds estimate the wavelengths to be 130 km and 113 km respectively. Because the propagation directions are so similar this picture first of all proofs that the wavelike behaviour at both times belong really to the same TID–event as had been exspected from the sonogram. After all this figure gives a good outlook of the possibilities one will have with further studies of the temporal evolution of propagating TIDs.

## AIMS OF FURTHER STUDY

As is mentioned earlier this paper is just a preliminary report of the study of atmospheric gravity waves and TIDs with the help of the EISCAT-CP2 data set. It explains only the principal methods of data preparation, of the resolution of different TID-events, and of the estimation of propagation directions. On the other hand this report shall also excite some interest on the further study of the observed atmospheric waves and TIDs with this excellent data base. These further aims will be:

- **Height dependence of the wave parameters**

The height resolution of about 20 km in the height range of 150 to 500 km offers the possibility to estimate the TID-events independently at all of these height intervalls. This means that one will get all wave parameters in the height range of 150 to 500 km independently at about every 20 km. Among the estimated parameters are especially the wave amplitude and the propagation direction, whose height and time–height dependence are of special interest.

- **The wave quantities in the different ionospheric parameters**

Because we are able to estimate the time and height dependence of the amplitudes and phases of the TID-events in each of the measured ionospheric parameters it is possible to study the phase differences of the same TID-event in the different ionospheric parameters. This offers totally new possiblities in the study of the relationship of the different ionospheric parameters with the wave in the neutral atmosphere and in the relationship of the ionospheric parameters among each other.

- **Special dispersion effects in the 'wings' of the sonograms**

The relatively good frequency resolution in the sonograms in the magnetically very quiet days allow further study of single TIDs which are just propagating through the measuring area. One can see these 'wings' i.e. one can see the temporal evolution of a propagating atmospheric gravity wave with respect to its period. It seems to be typical for the TIDs that the period first decreases a little and then increases slightly to periods being about 10-20% greater than the minimum period before they disappear. This must of course be studied further and has to be compared with theoretical considerations.

- **Generation of TIDs with the help of data from lower heights**

Beside the data from above 150 km height, EISCAT offer the same kind of data especially from about 60 to 150 km with a height resolution of only 3 km. These multipulse data from EISCAT allow exact estimations of the ionospheric densities and conductivities which together with the known electric field gives the ionospheric heating rates. With these data it should be possible to study exactly the generation of TIDs and especially estimate the amount of energy going into the waves.

- **Statistics for different magnetic activities**

Statistics of the data from different magnetic activity can be performed in order to develop some empirical relations between activity parameters and the amount of energy being stored, transported and released by large scale TIDs. Though it is not likely that there is only a global Kp dependence, it is perhaps possible to find special relations between magnetic activity indices (for example in Ae, Au or Al) and the occurence of local time dependent TID activity. This would then allow an improvement in the magnetic activity dependence of global atmospheric and thermospheric models.

## LITERATURE:


1. **Hines, C.O.**: *Internal atmospheric gravity waves at ionospheric heights, Can.J.Phys., 38, 1441-1481, 1960*

2. **Francis, S.H.**: *Global propagation of atmospheric gravity waves:A review J.atmos.terr.Phys., 37, 1011-1054, 1975*

3. **Bowman, G.G.**: *Movements of ionospheric irregularities and gravity waves, J.atmos.terr.Phys., 30, 721-734, 1968*

4. **Georges, T.M.**: *HF Doppler studies of travelling ionospheric disturbances, J.atmos.terr.Phys., 30, 735-746, 1968*

5. **Klostermeyer, J.**: *Gravity waves in the F-region, J.atmos.terr.Phys., 31, 25-45, 1969*

6. **Klostermeyer, J.**: *Influence of viscosity, thermal conduction, and ion drag on the propagation of atmospheric gravity waves in the thermosphere, Zeit. Geophys., 38, 881-890, 1972*

7. **Hooke, W.H.**: *Ionospheric irregularities produced by internal atmospheric gravity waves, J.atmos.terr.Phys., 30, 795-823, 1968*

8. **Volland, H.**: *The upper atmosphere as a multiple refractive medium for neutral air motions, J.atmos.terr.Phys., 31, 491-514, 1969a*

9. **Volland, H.**: *Full wave calculations of gravity wave propagation through the thermosphere, J.Geophys.Res., 74, 1786-1795, 1969b*

10. **Hines, C.O.**(*ed*): *The Upper Atmosphere in Motion, AGU, Washington, D.C., 1974*

11. **Beer, T.**: *Atmospheric Waves. Halsted, New York, 1974*

12 **Gossard, E.E. and W.H. Hooke:** *Waves in the Atmosphere, Elsevier, New York, 1975*

13. **Yeh, K.C. and C.H. Liu:** *Theory of Ionospheric Waves, Academic Press, New York and London, 1972*

14. **Richmond, A.D.**: *Gravity wave generation, propagation , and dissipation in the thermosphere, J.Geophys.Res., 83, 4131-4145, 1978*

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

16. **Williams, P.J.S., Crowley, G., Schlegel, K., Virdi, T.S., McCrea, I.W., Watkins, G., Wade, N., Hargreaves, J.K., Lachlan-Cope, T., Muller, H., Baldwin, J.E., Warner, P., van Eyken, A.P., Hapgood, M.A., Rodger, A.S.**: *The generation and propagation of atmospheric gravity waves observed during the worldwide atmospheric gravity-wave study (WAGS), J.atmos.terr.Phys., 50, 323-338, 1988*

17. **Lehtinen, M.S., Huuskonen, A.**: *The use of multipulse zero lag data to improve incoherent scatter radar power profile accuracy, J.atmos.terr.Phys., 48, 787-793, 1986*

18. **Richmond, A.D.**: *Thermospheric dynamics and electrodynamics; in Solar-Terrestrial Physics* by: Carovillano, R.L. and Forbes, J.M. (ed), *D.Reidel Publishing Company, Dordrecht, 1983*

19. **Schlegel, K.**: *The study of tides and gravity waves with the help of field-aligned velocities measured by EISCAT, J.atmos.terr.Phys., 48, 879-886, 1986*

20. **Ulrych, T.J., Bishop, T.N.**: *Maximum entropy spectral analysis and autoregressive decomposition, Rev. Geophys. Space Phys., 13, 183-200, 1975*

21. **Barrodale, I., Erickson, R.E.**: *Algorithms for least-squares linear prediction and maximum entropy spectral analysis - Part I: Theory, Geophysics, 45, 420-432, 1980*

22. **Barrodale, I., Erickson, R.E.**: *Algorithms for least-squares linear prediction and maximum entropy spectral analysis - Part II: Fortran program, Geophysics, 45, 433-446, 1980*

23. **Hines, C.O.**: *Propagation velocities and speeds in ionospheric waves: A review, J.atmos.terr.Phys., 36, 1179-1204, 1974*

## DISCUSSION

**T.Croft**

When TIDs are seen in sweep-frequency records of skywave backscatter at HF, they cause focussing on the distant Earth at a distance that varies with respect to radio frequency. This is attributed to a forward tilt of the TID. Since these TIDs must originate in a region and be initially non-tilted, it is expected that their tilt must develop and gradually increase as they travel. The rate of tilt development has never been measured, to my knowledge. I suggest that your data and your method of analysis may permit the rate to be observed, since your spectra show the age of each event and you can sample the tilt angle. This you may be able to derive tilt angle versus age.

(No reply requested - this is simply a suggestion.)

**B.Hunsucker**

There are about three or four papers published using the Chatanika results showing wave parameters in the source region. I will give you a list afterwards.

**Author's Reply**

Thank you.

**K.C.Yeh**

Since your observations are very close to the source region, are there occasions for which the arriving TID wave fronts are not plane?

**Author's Reply**

Up to now we have not estimated the wave fronts. The cross correlation of the (perhaps prefiltered) data of the 4 measured directions give us 6 time delays for every period we want to analyze. These time delays can be used for the diagrams of the phase trace speeds, but they can also be used to estimate the wavefronts. Ideally, we should then have 6 points of the wavefront so that we can contour it. In reality, there remain some problems; for example, the time length of the correlation interval in respect to the stationarity of the source or just the error in the estimation of the time delay which really produce an error bar instead of a point of the wavefront.

**K.Schlegel**

As the co-author of this paper, I would like to give some additional comments. First, Dr. Hunsucker mentioned that much of such investigations have already been done with the Chatanika radar. This is not quite true. EISCAT observes tristatically in 4 positions in the CP-2 measuring program. This is unique and enables us to determine the direction of the waves, which is not possible with any other radar.

Second, concerning the remark of Dr. Yeh, we mainly analyzed data from quiet days. In this case we can be sure that the observed waves are coming from a certain distance and that we are not close to the source region. On the other hand, on disturbed days we can calculate the source parameters, like Joule heating, particle heating or Lorentz-force, from the EISCAT data and take them properly into account.

**G.Rostoker**

Given the large wave activity throughout the interval where you identify specific gravity wave "pulses", how do you decide which pulse to select for analysis?

**Author's Reply**

We look at the systematically produced sonograms, which show the power of the variations at each period from 1 to 240 min. In very quiet times, it is possible to distinguish different "wave pulses" clearly as can be seen from the sonogram of the 5-6 May 1987. Till now we have not analyzed the more active periods, where the sonograms are more disturbed. In such cases one can look at the curves of the smoothed and properly filtered data at all heights. With such pictures of differently filtered data, it should be possible to decide if there are different "wave pulses" in that time period or not.

**I.McCrea**

Are there any instances in which you see variations in which the phase appears to be in phase at all heights?

**Author's Reply**

I have some plots where it's all in phase.

# ON THE POSSIBILITY OF PRODUCING ARTIFICIAL IONIZATION AND POLAR AURORA IN THE IONOSPHERE BY RADIO-WAVES EMITTED FROM THE GROUND

Mario Cutolo
*Angela Argenziano*
Department of Physics, University
of Bari (Italy), Centro Studi di
Radiopropagazione, University of Naples (Italy)

## Introduction

In 1938 Prof.V.A.Bailey (pioneer of the plasma physics) demonstrated by calculations that some resonance may occur between electrons and a wave as the frequency of this wave varies around the local gyrofrequency fh. A radio-wave having a frequency equal to the gyrofrequency is called by Bailey "gyro-wave". The resonance increases the collision frequency in the low Ionosphere by an appreciable amount (1).

There are two principal consequencies of the resonance:

1) If a second wave unmodulated passes through the ionosphere region where the "gyro-wave" acts, we have an interaction between these two waves. This phenomenon is called by Bailey "gyrointeraction" and has been demonstrated experimentally by Cutolo and colleaques (2).

2) If the radiowave is a gyro-wave, with a carrier frequency varying around the local gyrofrequency and is powerful it could produce in the Ionosphere an artificial aurora or airglow. The powerful gyro-wave would reduce strongly the electric voltage of the region illuminated by the gyro-wave as to generate an alectrical discharge. Moreover Bailey has shown that with a suitable system of Beam antenna and with a power of 500 Kw or of 1.000.000 Kw it would be possible to generate an artificial aurora between 60 and 88 Km (3) in the Ionosphere. The artificial aurora would consist, with a diameter of 9 Km and height 20 Km or in the case of a power of 1.000.000 Km the aurora would appear as a luminous sphere with a diameter 9 times greater than that of the full moon but with minor brightness (1).

The aurora would appear between 65 and 88 Km from the surface of the earth.

V.A. Bailey starting from the experiments of H.A. Wilson and G.M. Myers made in Sydney with high-frequency discharges on the lines developed by J.S. Towsend and his associates developed his theory.

These experiments made Wilson and Myers showed that the electric discharges can occur in the air when in the uniform column z/p the ratio of the electric force z to the gas pressure p is approximately 16. It is possible therefore to conclude that two of the necessary conditions for producing electrons and light are satisfied when E/p 16.

A new investigation was made in 1959 by V.A. Bailey of the possible increases of both the collision frequencies and the electron density N caused by a powerful extraordinary circular gyrowave in the nocturnal lower-F-region and in the daytime D-region.

The theory was based on the most reliable available experimental data on the behaviour of free electrons in the air with energies up to 1.12 eV.

The principal process then freshly introduced was the attachment of the electrons to molecules in some collisions. According to his new theory the power density requirement for two excitations of an artificial airglow by means of gyrowaves were reconsidered by Bailey in 1961 in the light of later information. The important conclusion was that an array of 80 dipoles radiating vertically a beam of gyrowaves with a mean power of 500 Kw would produce an easily observable enhancement of the natural airglow.

## I. New technique to generate an artificial aurora

A long series of experiments of the gyrointeraction phenomenon was in Italy from 1946 to 1950 and as a result we found the resonance curve with two humps (bactrian curve) and the resonance curve with one hump (dromedary curve) (4).

During the experiments of 1947 the writer observed that the resonance increased when the Radio-transmission was carried out by pulses method instead of cm (5). This was due to transient actions and not to increased power because we were not able, for technical reasons, to increase the anodic voltage of the Final Amplifier of the transmitter. Later we tried to generate an artificial aurora in the Ionosphere, Following the suggestion of V.A. Bailey, by means of an intense emission of Radiowaves with carrier frequency equal to the local gyrofrequency.

In order to make a detailed study we tried to obtain the phenomenon of aurora through laboratory experiments. To measure an eventual airglow in the Ionosphere, I thought of reproducing the phenomenon in the laboratory using a glass globe containing dry air at same pressure as the Ionosphere.

I also thought of using a radiative electromagnetic field instead of an inductive field as other researchers do. To use the radiative field it is necessary that the oscillator producing the radiowaves be

more than a wave length away from the glass-globe.

To carry out the experiment a transmitting and a receving system are necessary. Many experiments including a remarkable one made on 8 Novembre 1956 were carried out.

The transmitter emitted pulse waves of 64 MHz with mean power of 70-80 Watt and peak power of 1.8 Kw.

The antenna system had a gain of 16 dB. The transmitter was installed on the terrace of the Technical Physics Institute of the University of Naples (Italy). The receiving equipment was located in another building at a distance of 1.4 kilometres as the crow flies from the transmitting place.

It consisted of a half wave antenna (located on the terrace of the receving building) connected with a cable 20 m long, that brought the radiofrequency from dipole to the Lecher line. The other end of this line had a condenser consisting of two metallic parallel plates distant 19 cm one from the other.

A glass-globe containing dry air at low pressure ($2x10^{-3}$ mmHg) was placed between the armatures of the condenser.

According to Larmor equation for F=64 MHz, H=22.80 gauss. When the transmitter (located at 1.4 Km) emitted pulse waves and the magnetic field began to work, sensible luminescence was generated in the globe.

The discharge voltage capable of generating the visible luminescence was $2.10^{-2}$ Volt/cm, as $p=2x10^{-3}$ mmHg we have that $E/p = 2x10^{-2}/2x10^{-3} \approx 10$.

After this result I decided to try to generate the artificial airglow (aurora) in the Ionosphere.

At first I thought of sending towards the Ionosphere an intense pulse beam of Radiowaves with carrier frequency almost equal to the local gyrofrequency. A long report of this plan was shown to the General Staff of the Italian Defence Ministry. During long discussion some perplexity arose among the Commission members. According to the general opinion the possibility of concentrating an intense beam of Radiowaves in a small zone of the Ionosphere is difficult to realise with medium frequency of 1 or 2 MHz. in fact the gyrofrequency at 45° latitude is 1.2 MHz.

For this a wave length of 200 metres, that is too high.

For this I thought of using V.U.F. or U.H.F. waves, but modulate with frequencies varying around the local gyrofrequency.

Without modulation with frequency equal to the local gyrofrequency the electric field of the V.H.F. or U.H.F. waves would be too high. In fact V.A. Bailey showed in 1938 that the action of the terrestrial magnetic field is to lower greatly the discharge voltage.

As early as 1953 it was foreseen that owing to the clearly non-linear behaviour of D region, particularly of Bailey's layer (60-88 Km) the plasma may act as a Detector, separating the low frequency from the carrier of the modulated wave (Cutolo 1964) (6).

Consequently it is to be expected that if a VHF or UHF wave, modulated with a frequency equal to the local gyrofrequency, is radiated towards the Ionosphere, the wave envelope would have some effect on the plasma.

To verify this prediction, in 1957 a suitable technique was devised, described in a patent applied for in the United States at the end of 1957 and later granted (7).

Experiments conducted with VHF waves modulated at the gyrofrequency, in Naples from 1953-1962 confirmed the predictions-that is, the separation of the low frequency from the carrier modulated with a frequency equal to the local gyrofrequency (1.2 MHz).

We made many experiments to verify the existence of the phenomenon called "Detection effect". The most important were made during the nights of March 3, 5, 7 and 9 1962 from 0100 to 0400 T.E.M.E.C. In order to confirm that existence, we decided to use the gyrointeraction phenomenon.

The Augusta wave of the Italian Navy (540 KHz) worked as a wanted station.

The V.H.F. transmitter (50-75 MHz, 80 Kw) acted as a perturbing station. The modulation frequency varied from 1100 to 1300 KHz to pass through the local gyrofrequency (1200 KHz). The peak value of the power radiated was 80 Kw and the modulations depth was 40-45 percent. The aerial system was of the Yagi type having a theoretical gain of 10-12 of dB and irradiating the Ionosphere 90 Km above Sessa Aurunca half way between Augusta and Venice.

The receiving station with receiver and measurements equipment was located in Venice.

During the nights of 5th and 9th the effect was recorded with the cross-modulation percentages exceeding 20%.

Many other experiments were made at vertical incidence using as wanted a station located at Nola and as receiving station placed at Brusciano. The V.U.F. transmitter was located in Naples.

While the experiments conducted with oblique incidence established the existence of an Ionospheric Detection effect, those carried out with almost vertical incidence showed that the phenomenon of the separation is rather weak according to Lanza's theory.

Interpretations of the phenomenon were given by Bailey, Caldirola, Gurevich (8) Merzel (9) and Lanza. Unlike Bailey and Caldirola the theory that follows Cutolo's ideas, i.e., admits a Detection effect in the radioengineering sense, is the one proposed by Lanza (10) of the North Eastern University of Boston.

All these theories are based on "Non-linearity of the D region or noctume E layer.

Since the number of electrons is very small in these parts of the Ionosphere (at night N 100 electrons/$cm^{-3}$ or even less) and the mean free path is considerable, between one collision and another with the neutral molecules the electrons have time to acquire a fair amount of energy from the electric field of the radiowaves crossing the D region of the Ionosphere.

But since the mass of the electron is very small compared to that of the molecules the electron does not transfer to the latter all the energy acquired from the electric field so that it heats up, with the result that the dielectric permittivity of the plasma begins to depend on the electric field E.

It follows that the polarization and the conduction current are no longer proportional to E, so that all the electro-dynamic processes, in particular the propagation od radio waves, acquire the property of no longer being linear, with the consequent violation of the principle of superposition.

If the plasma is not linear, it is clear that it may act as a Detector or a Transistor or a Diode.

Another consideration is the following. If we transmit a wave with carrier frequency equal to the gyrofrequency this wave would be absorbed while crossing the D-region reaches the wanted wave too weak. If we instead use a V.H.F. or an U.H.F., both modulated at gyrofrequency and the separation occurs in the Ionosphere (because this last is non linear and acts as a low-pass device) the gyrofrequency field liberated from the carrier by detection can reach any height without be absorbed. The separation of the gyrofrequency from the carrier permits it to act on the Ionosphere increasing the collision frequency.

If the increasing of the collision frequency is notable, we can have the production of fresh electrons and light.

If the experiments of Detection effect were made at high latitude, the phenomenon would be more intense because at that latitude the Resonance effect would be more intense because the angle between the electric field of the Radiowaves V.H.F. and U.H.F. and the Terrestrial Magnetic Field is nearly 90°.

To establish further the existence and the mechanism of the Detection effect (that is the separation of the modulating frequency from the carrier), we made some experiments at Arecibo Observatory (USA) in 1976 (11).

The transmitter emitted at 430 MHz, modulated with frequencies varying from 1020 KHz to 1156 MHz, to pass through the local gyrofrequency. A Collins receiver equipped with advanced technique and computer was tuned to the modulation frequencies.

The presence of signals, with frequencies varying around the local gyro-frequency recorded by the computer, shows that the separation of the modulating wave (the gyro-frequency) from the carrier (430 MHz) takes place. As it might be shown from experiments, the behaviour of the amplitude of the signals is exponential. Moreover the amplitude of the signals is very little because they are very weak and this is in accordance to the theory.

In fact, after the separation of the modulating frequency from the carrier, the signals are scattered in all directions more easily than in the vertical direction.

The greatest quantity of the gyro wave, after the separation, is naturally highly absorbed from the electrons of the plasma for resonance. The remarkable energy acquired from the electrons strongly increases their speed so that they collide violently with the molecules producing fresh electrons and airglow.

The artificial production of fresh electrons permits the propagation of radiowaves of higher frequency through the Ionosphere and therefore U.H.F. radiocommunication at great distances.

The production of fresh electrons could be useful for the defence of a country.

If the experiments of Detection effect were made at high latitude the phenomenon would be more intense because at that latitude the resonance effect would be stronger.

The angle between the electric field of the radiowaves V.H.F. and U.H.F. and the terrestrial magnetic field is near 90°.

It has been demonstrated that the work done by the electric field of a radiowave on an electron between two consecutive collisions with molecules is given by

$$W = \frac{e^2}{2m} E_e^2 \left( A + A \cos 2\omega t + B \sin \omega t \right)$$

where Ee is the actual value R.M.S. of the electric field E

$$A = \frac{sen^2\Psi}{\nu^2+(\omega-\Omega)^2} + \frac{sen^2\Psi}{\nu^2+(\omega-\Omega)^2} + \frac{2\cos^2\Psi}{\nu^2+\omega^2}$$

$$B = -i\left(\frac{\omega-\Omega}{\nu^2+(\omega-\Omega)^2} + \frac{\omega+\Omega}{\nu^2+(\omega+\Omega)^2} - \frac{sen^2\Psi}{\nu} - \frac{2\omega}{\nu^2+\omega^2}\right)$$

$\omega$ = the angular frequency of the wave

$\Omega$ = eh/ mc angular gyro-frequency

$\Psi$ = angle between E and H (terrestrial magnetic field)

$\nu$ = the mean value of the collision frequency of an electron with the neutral molecules of the plasma.

From here we see that the quantity A has maximum value when the angle is 90° and that reaches a large amount not only for $\omega = \Omega$ but also for $\Psi$ =90°. In this case the resonance has the maximum intensity and hence the electrical discharge starts more easily.

ACKNOWLEDGEMENT

The author wishes to thank his colleague Dr.V.Cutillo, Prof.C.Goutelard, Prof.L.Bossy, Dr.Col.P.A. Brunelli for their interest to my paper and Miss M.McLean for the careful revision of the English and Miss F. Gaglione for typing the manuscript.

REFERENCES

1) Bailey V.A. Phil.Mag.26, 425, 1938
Bailey V.A. , R.A. Smith etc., Nature, 169, 911-913, 1952
2) Cutolo M., Carlevaro M., Gherghi M.,Alta Frequenza, June 1946
3) Smith P.A. and collaborators, Report N.3 and Final Report, University of New England [illegible] Conference Proceedings, 11-15 April 1966, Ottawa, Canada, Vol.I; Dicembre 1967
4) Cutolo M., G.Agnelli, M.Cimino and Puglisi , Advances in Physics, Advances in Physics, 14, [illegible]
5) Cutolo M., Nature 160, 834, 1947
6) Cutolo M. AGARDgraph 74, 35-42 1964
7) Cutolo M., United States Patent Office, June 15, 1965
8) Gurevich A.V., Rivista di Fisica Teorica, URSS, 1968
9) Menzel D.H., Radio Science, 69D, 1-7, 1965
10) Lanza G., Conference Proceedings, 11-15 Aprile 1966, Ottawa Canada Vol.I, p.36[illegible], Dicembre 1967
11) Cutolo M. G.Colonnese: AGARD Conference Preprint N.263, 1979

## DISCUSSION

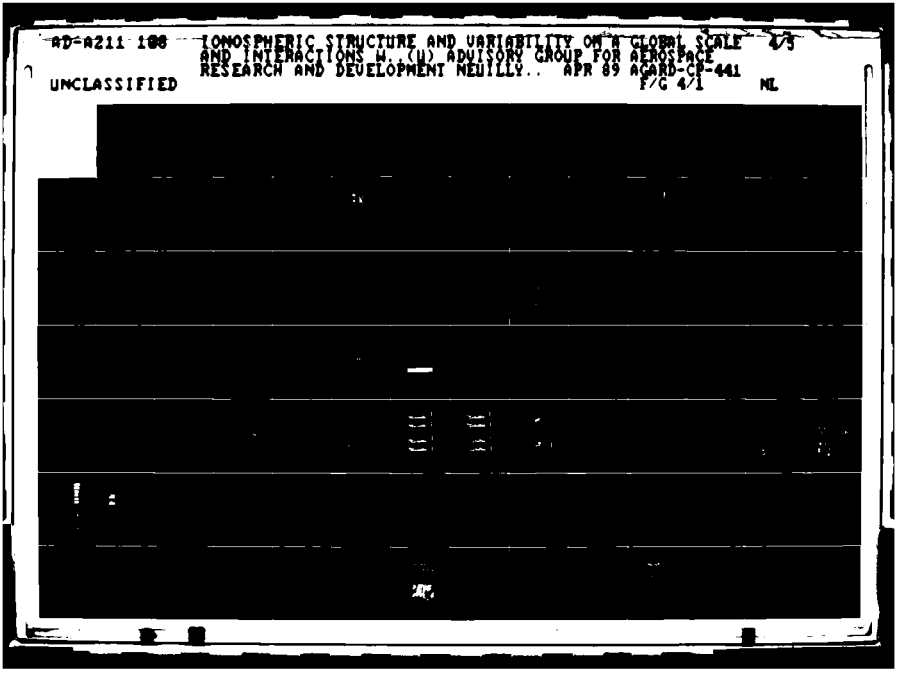
**R.Showen**

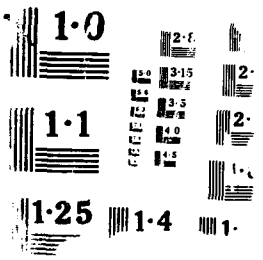
Has the work you did at Arecibo been published and has anyone else been able to independently verify what you have said?

**Author's Reply**

The work I did at Arecibo has not yet been published.

AD-A211 188 IONOSPHERIC STRUCTURE AND VARIABILITY ON A GLOBAL SCALE 4/5
AND INTERACTIONS W.. (U) ADVISORY GROUP FOR AEROSPACE
RESEARCH AND DEVELOPMENT NEUILLY.. APR 89 AGARD-CP-441
UNCLASSIFIED F/G 4/1 NL




1·0
1·1
1·25
1·4
1·

# IONOSPHERIC IRREGULARITIES DUE TO POWERFUL HF RADIO TRANSMISSIONS

by
J. A. Fejer
Department of Electrical and Computer Engineering
University of California, San Diego, C-014
La Jolla, California 92093
USA

## SUMMARY

It has been known for some time that sufficiently powerful HF radio transmissions produce a great variety of ionospheric irregularities. Ionospheric modification facilities with effective radiated powers of the order of 100 MW directed toward the ionosphere have been used for the study of such irregularities since 1970. Such man-made irregularities have been employed to establish experimental scatter communications links. Very powerful short wave broadcast transmitters must also produce man-made irregularities which affect the ionospheric propagation of short waves.

All aspects of the physical phenomena which play a role in the production of ionospheric irregularities by powerful HF transmissions, are discussed. These include thermal self-focusing of radio waves, formation of short-scale field-aligned irregularities by a thermal parametric instability in which the scattering of the HF pump wave by the irregularities into Langmuir waves plays an important role, and those parametric instabilities in which the ponderomotive force dominates over thermal forces. The latter two parametric instabilities can lead to the acceleration of electrons to energies of tens of electron volts. Such accelerated electrons can produce artificial airglow and also additional ionization which under certain conditions could be significant. In their strongly nonlinear stage parametric instabilities can lead to the formation of localized electron density depletions (cavitons) maintained by the ponderomotive force of Langmuir waves trapped in them.

## 1. INTRODUCTION

Powerful HF radio transmissions radiated by modern ionospheric modification facilities produce irregularities by a process of stimulated scattering. The irregularities scatter the radio wave into other high frequency waves which could be electrostatic (Langmuir waves) or electromagnetic (other radio waves). These other high frequency waves interact with the original powerful radio wave via a nonlinear force which by feedback causes the irregularities to grow (stimulates them), starting from their initial thermal level. The growth of the irregularities is thus due to a type of plasma instability which goes by the names of either stimulated scattering or parametric instability.

The irregularities could be independent of time or they could change "slowly", the corresponding frequencies being very much smaller than that of the original powerful radio wave. In this sense an assembly of ion acoustic waves will also be referred to as irregularities. It is in this more general sense that ionospheric irregularities due to powerful HF radio transmissions will be discussed here.

It is usually easiest to consider first the linear theory of a parametric instability in a homogeneous medium; in practice even an inhomogeneous medium can often be approximated locally by a homogeneous medium. Such a theory considers the initial stage when the growing irregularities are of small amplitude; then their growth and their possibly oscillatory nature can be treated by Fourier analysis. In many cases a wavelike Fourier component could have zero frequency and then the irregularity is simply a spatially periodic density perturbation of small amplitude. A linear mathematical formulation of the stimulated scattering process described above then leads to a dispersion relation which for a real wave vector $\underline{k}$ of the assumed wavelike irregularity leads to a complex angular frequency $\omega$.

The dispersion relation depends on the strength of the original powerful radio wave, usually called the pump wave. For a sufficiently weak pump wave all the solutions of the dispersion relation are damped waves if the plasma is stable. As the strength of the pump wave is increased, eventually for a certain wave vector $\underline{k}$ the dispersion relation leads to a real angular frequency; the condition for this is called the linear threshold condition. When the strength of the pump wave is above its threshold strength then the linear dispersion relation results in a positive growth rate for a range of values of the wave vector $\underline{k}$ of the assumed small irregularity. The growth rates thus obtained are called linear growth rates and are only valid for small amplitudes.

Linear theory gives no information on the mechanism that eventually stops the growth and results in the so called saturation spectrum of the irregularities. The nonlinear theories of the saturation of parametric instabilities are less well developed than the linear theories on account of their more difficult nature but some progress has been made.

Each of the following sections will deal with one of the different parametric instabilities excited by a powerful radio wave. Their associated low or zero frequency

irregularities and the high frequency waves into which the irregularities scatter the pump wave will be discussed from both a theoretical and an observational point of view. Although in reality the different parametric instabilities are not excited completely independently of each other, it will be convenient to deal with them one at a time without ignoring their interdependence.

## 2. THE OSCILLATING TWO STREAM INSTABILITY

This parametric instability [1]-[3], usually abbreviated OTSI, was the first one whose excitation in ionospheric modification experiments was predicted theoretically [4]. In this instability the electromagnetic pump wave of ordinary polarization causes the simultaneous growth of a spatially periodic non-oscillatory density perturbation and a standing Langmuir wave of about the same wavelength, both being very much smaller than the wavelength of the pump. Since the Langmuir wave is the result of scattering of the pump wave by the density perturbation, the frequency of the Langmuir wave is equal to the frequency of the pump. In the linear theory this instability and the parametric decay instability, discussed in the next section, represent two solutions of the same dispersion relation. A brief outline of the theory will therefore be postponed; it will form part of the next section. It should be remarked already here, however, that these two instabilities can only be excited by a pump wave of ordinary polarization at frequencies below the maximum plasma frequency of the ionosphere; the extraordinary wave is reflected below the height where lightly damped Langmuir waves can exist. In both instabilities the nonlinear force completing the feedback loop is the ponderomotive force.

The OTSI has been detected at Arecibo [5], [6] and later at Tromso [7] by observing the upshifted or downshifted spectra of radar backscatter by its Langmuir waves. In those observations a radar pulse whose length was of the order of 1 ms was used. The radar echo was then coherently sampled, say, every 2 microseconds. The complex digital Fourier transform of the sample values was then taken and the absolute values squared to obtain the power spectrum with a frequency resolution of about 1 kHz, the reciprocal of the pulse length of 1ms. Such spectra, which in addition to the OTSI also show the excitation of the PDI (parametric decay instability), will be shown in the next section; they are usually called spectra of the enhanced plasma line,to differentiate them from the plasma line of incoherent scatter.

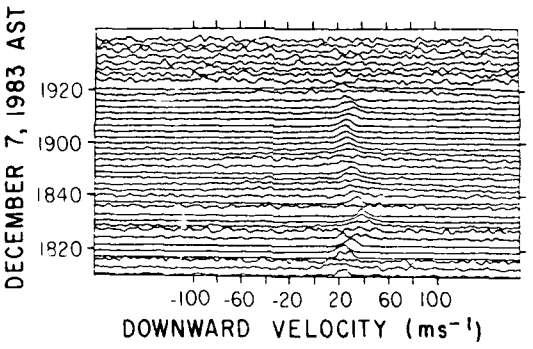



Figure 1. Spectra of the OTSI line over a bandwidth of 1 kHz, calibrated in line-of-sight drift velocity, for different local times.

In this section 1 kHz wide spectra, obtained by using much shorter radar pulses transmitted coherently every millisecond and by sampling a large number of consecutive scatter echoes, will be shown. Such spectra have a much better frequency resolution, determined by the reciprocal of the the duration of the sequence of short pulses. Figure 1 (Figure 7 of [8])shows such 1 kHz wide spectra taken every minute for about an hour As will be seen in the next section, the spectra of the PDI are very much wider than 1kHz and appear aliased as noise in the spectra of Figure 1. Those spectra therefore only display the (upshifted) radar backscatter from the Langmuir waves of the OTSI which in the absence of ionospheric drifts would have a frequency of 435.1 MHz, the sum of the radar frequency of 430 MHz and the HF transmitter frequency of 5.1 MHz. The frequency of the Langmuir waves of the OTSI are predicted to be equal to the frequency of the HF pump wave in a reference frame drifting with the electron gas; the Doppler shift due to that drift velocity causes the frequency of the upshifted OTSI line to differ from 535.1 MHz, marked 0 m/s on Figure 1 whose 1kHz wide spectra are calibrated in line-of-sight drift velocity. The maximum plasma frequency of the ionosphere sank below the pump frequency of 5.1 MHz at about 1920 AST and the spectra obtained after that time do not show the OTSI scatter echo.

Figure 1 shows that the line-of-sight electron drift velocity may be determined from observations of the OTSI. Observations of the ion line of incoherent scatter are known to yield the line-of-sight ion drift velocity. It should therefore be possible

to determine the line of sight component of the current density vector from the combination of those two observations.

## 3. THE PARAMETRIC DECAY INSTABILITY

In the PDI the generated low frequency wave is an ion acoustic wave and the generated high frequency wave is a Langmuir wave whose frequency is less than the pump frequency by the frequency of the ion acoustic wave. This follows from the matching relations

$$\omega_0 = \omega_1 + \omega_2 \qquad (1)$$

$$\underline{k}_0 = \underline{k}_1 + \underline{k}_2 \qquad (2)$$

satisfied by the pump wave and the two generated (daughter) waves in parametric instabilities of this type.

The dispersion relation resulting from the linear theory was derived, for example, in [9]. The feedback to the low frequency wave results in this case from the ponderomotive force. The growth rate resulting from that dispersion relation is largest when the propagation vector is parallel to the pump electric field. For that direction the growth rate has two well separated maxima as a function of the wave number, corresponding to the PDI and the OTSI. Although the threshold pump electric field is slightly lower for the PDI than for the OTSI, both being slightly below 0.2 V/m , Figures 1 and 2 of [10], based on computations by Erhan Kudeki, show that for typical ionospheric parameters the linear growth rate of the OTSI exceeds that of the PDI for pump electric fields greater than 0.5 V/m when X = 0.983, and for pump electric fields greater than 0.9 V/m when X = 0.997 where X is the square of the ratio of the plasma frequency to the pump frequency. Thus for full pump power of a typical ionospheric modification facility the linear growth rate of the OTSI certainly exceeds that of the PDI at the first peak of the Airy pattern but probably not at the lower height where Langmuir waves detected by the 430 MHz radar at Arecibo satisfy the dispersion relation in the unpertubed ionosphere.

Saturation spectra predicted by weak turbulence theory were computed in [11] and [12]. The results of those computations show that the Langmuir waves of the parametric decay instability, excited first linearly by the pump, eventually decay further parametrically into Langmuir waves going in the opposite direction. Several additional similar decays produce a whole cascade of Langmuir waves, the frequency difference between the pump and the first member of the cascade being half of the frequency difference between neighboring members of the cascade. The propagation vectors of the excited Langmuir waves form angles of less than about 25° with the pump electric field which for the ordinary wave is close to the direction of the geomagnetic field near the height of reflection.

The Arecibo radar beam forms an angle of 45° with the geomagnetic field and therefore according to theory the radar should not detect the parametrically excited Langmuir waves directly. It should detect ([12], [13]), however, the much weaker Langmuir waves into which the parametrically excited Langmuir waves and the pump wave are scattered by the ion acoustic waves in thermal equilibrium (which are also responsible for incoherent scatter). Even the Langmuir waves into which the pump wave slightly below threshold strength is scattered by the thermal ion acoustic waves, are still well above the thermal level of Langmuir waves [14].

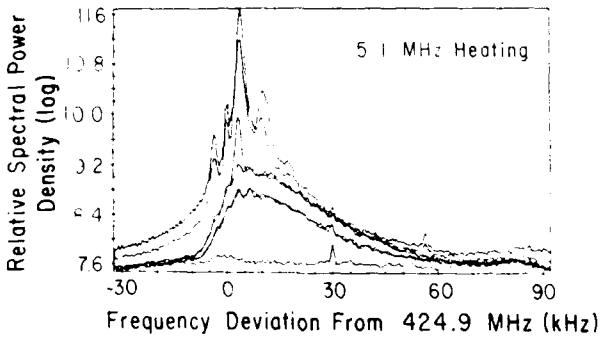



Figure 2. Several spectra of the enhanced plasma line obtained by scanning the angle of the 430 MHz radar beam while keeping the elevation angle constant at 3.5° from vertical. All the spectra were obtained within a single scan.

Figure 2 (Figure 9 of [8]) shows some downshifted enhanced plasma line spectra from the F-region above Arecibo, ranging from below threshold conditions to full pump strength.

The weakest spectrum, below threshold, is below receiver noise level by a factor of about 4 and is double-humped; a better example of the same type of spectrum is shown by Figure 3 of [8].

The two next weakest and almost identical spectra of Figure 2 are not too different from those predicted in [13] and shown by the strongest spectrum of their Figure 6 which incidentally also shows a spectrum below threshold.

The two strongest spectra of Figure 2 are the spectra most commonly observed with CW operation at full transmitter power, typically 80 MW equivalent radiated power (ERP). The peak at 424.9 MHz, marked 0 on the abscissa axis, is believed to be due to the excitation of the OTSI. The much stronger peak, called the decay line in Arecibo parlance, at 424.9MHz + 3.6kHz (representing Langmuir waves near frequencies of 5.1 MHz - 3.6 kHz) is believed to be the first member of the PDI cascade; the frequency difference of 3.6 kHz is the value obtained from the dispersion relation for the highest linear growth rate of the PDI, for a temperature of 800 K and a wavelength of 35 cm. Note that this peak is 1000 times above receiver noise and nearly 40 times above the maxima of the strongest spectra still resembling those predicted by Figure 6 of [13].

In addition to the strong peak at 3.6 kHz, there are weaker peaks at about 10.8 kHz and 18kHz which could be interpreted as further members of the cascade, spaced 7.2 kHz apart. The very weak peak at 424,9 MHz - 3.6 kHz is the result of scattering by the 3.6 kHz ion acoustic wave of the 5.1 MHz pump wave into a 5.1 MHz + 3kHz Langmuir wave. In contrast to the decay line, this peak is the result of non-resonant scattering.

The two strongest spectra of the type shown by Figure 2 do not agree with the theories of Fejer and Kuo (1973a, b) and Perkins et al. (1974). First of all those theories predict that the Arecibo radar should not detect directly the Langmuir waves of the cascade but only those much weaker Langmuir waves resulting from the scattering of the pump and of the waves of the cascade by the thermal level of ion acoustic waves. Secondly the observed power ratios of the different members of the cascade are very much greater than predicted for the Langmuir waves of the cascade.

Qualitative attempts [8] have been made to explain the discrepancy in terms of the short-scale field-aligned irregularities discussed in the next section. However, since those attempts, further observations resulted in even greater difficulties in the way of a theoretical explanation of the strongest spectra of Figure 2.

Simultaneous observations of the enhanced and the natural plasma line ([15], [16]), using the technique of chirped radar pulses [17], confirmed an earlier suggestion based on observations of a different type [18], that for high pump powers the enhanced plasma line is generated near the reflection height of the ordinary wave and not at the lower height where Langmuir waves detected by the 430 MHz radar satisfy the dispersion relation in the unperturbed medium. Moreover, the observations in [16] have been interpreted in terms of localized density depletions filled with Langmuir waves whose ponderomotive force causes the depletions (cavitons) which are formed in less than 10 ms after switching on the pump and which are thought to have linear dimensions of the order of 10 m. For details of the technique of the observations and of the justification of the above interpretation the reader is referred to the cited publications.

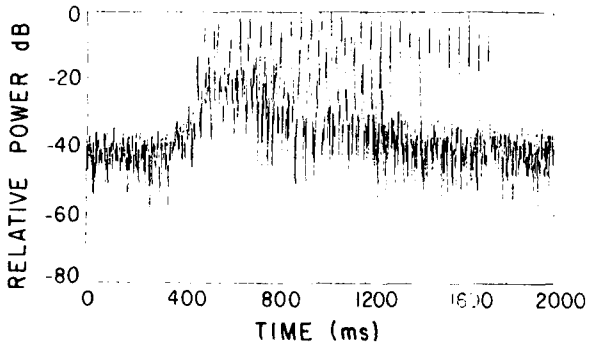



Figure 3. Quasi-periodically recurring pulses in the relative power in db of the enhanced plasma line with a 2s on 2s off transmitting sequence, sampling the power every millisecond. A 2s long period starting at 163209 on 19 January, 1987 is displayed.

Another recent set of observations [19], following up earlier suggestions [8], indicates that the temporal behavior of the power in the enhanced plasma line at Arecibo is much more complicated than was previously believed. In an extreme case, shown by Figure 3, the power consists of pulses of about 5ms duration, repeating regularly

at intervals of 42ms. This is seen especially clearly between the 1200 ms and the 1700 ms marks. The pump wave was switched on near the 500 ms mark as part of a 2 s on 2 s off cycling process of long duration. Note that the pump remains on until 2500 ms but the power in the enhanced plasma line drops to zero after about 1700 ms in spite of the continued presence of the pump. This "overshoot" phenomenon will be discussed in the next section.

Although this extreme type of behavior is seen only over a relatively small percentage of the time, the temporal variation of the power in the enhanced plasma line is unlike that of random noise over a large percentage of the time. Behavior similar to that of random noise would be expected for a stationary random process.

There is probably a close relation between the the finite bandwidth of the OTSI spectra of Figure 1 and the temporal behavior of the enhanced plasma line power, especially when it is realized that the OTSI is generated at the same height as the decay line and its cascade [19], near the reflection height of the pump [16]. Moreover, a strong positive correlation was found between the bandwidth of the OTSI and the pump power [19]. This suggests that the variations in the power of the enhanced plasma line are speeded up as the pump power is increased. Clearly more observations and further development in the theory are needed to understand the physical processes involved in the generation of the enhanced plasma line.

Another puzzling aspect of the problem, not mentioned so far, is the complete quantitative disagreement of the Arecibo observations [19]) with numerical simulations of Langmuir cavitons at Los Alamos [20], in spite of the superficial similarity between their sequence of bursts of Langmuir oscillations in their caviton simulations and the pulses shown by the present Figure 3.

## 4. THERMAL PARAMETRIC INSTABILITY

This instability was studied extensively by the ionospheric modification facility at Platteville [21] by aspect sensitive radar backscatter from its field-aligned irregularities [22] and by forward scatter [23]. The low frequency daughter "waves" of this instability are short scale (0.3 - 20 m) field-aligned irregularities and the high frequency waves are Langmuir waves. The feedback is in this case due to a thermal, partial pressure force which dominates over the ponderomotive force. The condition for this dominance [24] is that the length of the irregularity should be greater than the electron mean free path. This condition is certainly satisfied in the F region where the electron mean free path is about a kilometer.

Barry [23] demonstrated the limited applicability of these irregularities to communication by establishing a temporary scatter communication link between the West Coast in California and the East coast in Texas, by means of artificially produced irregularities over Platteville, Colorado.

The generation of similar short-scale field-aligned irregularities over the modification facilities near Gorki [25], Tromso [26], and Arecibo [27] was also demonstrated.

Theories of this instability were formulated [28]-[34] some years after the first observations. Essentially all these theories consider a modulational instability which differs from the OTSI in that the nonlinear force responsible for the feedback to the irregularity is not ponderomotive but thermal in nature, and in that the inhomogeneous nature of the ionosphere plays a more important role [31], [32]. In fact the inhomogeneous nature of the ionosphere prevents a linear instability unless the standing wave nature of the pump wave is taken into account [31], [32] although nonlinear theory shows that growth occurs if the initial density perturbation exceeds a critical amount [29], [30], [33]. This critical amount could be provided by the mechanism of [31], [32] or by the mechanism of [28]. The latter was the first theory of this instability; it used the Langmuir waves of the parametric instability acting as pump. Inhester [34] carried this instability furthest into the nonlinear stage in his simulations and showed how the irregularities become narrower and more localized in a direction perpendicular to the geomagnetic field as they grow.

It should be noted that the Langmuir waves of this instability propagate, at least initially, perpendicularly to the geomagnetic field. These Langmuir waves are therefore excited below the height where the pump frequency equals the local upper hybrid frequency, well below the greater height where the pump wave is reflected and the pump frequency equals the local plasma frequency.

It should also be noted that the instability is excited by the perpendicular component of the pump electric field at a height slightly below where the pump frequency is equal to the upper hybrid frequency. That component strongly depends on the angle of inclination and has a much smaller amplitude at Arecibo than at Tromso. Observations (with the radars at Guadeloupe [27], Lycksele [26] and White Sands [21], [22]) confirm that the backscatter from short-scale field aligned irregularities is very much weaker at Arecibo than at Tromso or Platteville.

Short-scale field-aligned irregularities are believed to play a role in the overshoot phenomenon of the enhanced plasma line [35] whose main feature is a drop to a lower or zero level in the power in the enhanced plasma line during a time of the order of a second after having reached maximum level in a time much shorter than a second

after the pump is switched on. An example of the overshoot is shown by Figure 3 with a somewhat better temporal resolution than in the observations of [35]. Graham and Fejer [36] suggested that the overshoot is caused by the absorption suffered by the pump wave near the height where its frequency is close to the upper hybrid frequency and the field aligned irregularities scatter the pump wave into Langmuir waves. Fejer and Kopka [37] demonstrated the absorption of the reflected pump wave of the Tromso heating facility, both by this mechanism and by the excitation of the parametric decay instability; however Djuth and Gonzales [38] showed that at Arecibo such absorption is sometimes too small to be observed and can not explain the observed overshoot. Other Arecibo observations [19] suggest that even on such occasions the growth of a plasma instability, presumably that responsible for the short-scale field-aligned irregularities, is the cause of the overshoot.

It should be noted that both the thermal parametric instability discussed in this section and the two striction type parametric instabilities, the PDI and the OTSI, discussed in the last section, generate Langmuir waves which can accelerate electrons. Collisional excitation of the 630 nm and 557.7 nm oxygen lines by such electrons was observed very early [39]). The question arises whether electron acceleration is mainly due to the Langmuir waves of the thermal parametric instability or of the striction parametric instabilities. Workers in the USSR favor the thermal parametric instability [40]. Fejer and Sulzer [41] observed the accelerated electrons by observing the enhancement in the natural plasma line caused by them at night. They show that electron acceleration is strongest before the growth of the thermal parametric instability and must therefore be caused by the Langmuir waves of the striction parametric instability. It seems likely that the striction parametric instabilities play a larger role in electron acceleration at lower latitudes than the thermal parametric instability but the latter may well be more important at higher latitudes with the exception of the first second or so after the heating transmissions are switched on.

## 5. THERMAL SELF-FOCUSING

This instability was one of the first physical processes discovered when the Platteville facility started operation [42]. Within seconds of switching on the modifying HF transmissions the ionograms showed the presence of spread-F although the full extent of spread-F took some minutes to develop. The irregularities produced by the instability were also investigated by the observation of enhanced scintillation of signals observed from satellite transmitters [43], [44] and from radio stars [45]-[48]). Another method of investigating the irregularities is the observation of the fading of the power in the enhanced plasma line on time scales of the order of 10 s [49]-[51]. The density variations were also measured directly by satellite in situ probes [51]. These observations show the presence of artificial ionospheric irregularities over a scale length range of a few kilometers to tens of meters. Figure 4 (Figure 4 of [51]) shows the density perturbations measured by the satellite and also illustrates its path through the estimated heater beam.

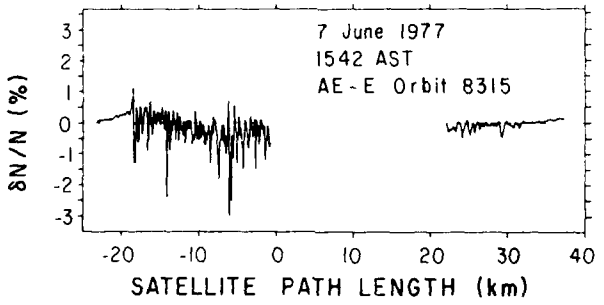



Figure 4. Detrended satellite observations of density fluctuations in the heated region. The origin is the point along the orbital path which is due north of Arecibo. The data gap results from the fact that the satellite operates 3 s on 3 s off in this mode.

The basic idea of thermal self-focusing is rather simple. Initial small plasma density perturbations cause focusing and defocusing of the heating wave. The focusing in the density depletions causes enhanced heating and further depletion, allowing the perturbations to grow. The theories are in a sense simpler when the heating wave penetrates the ionosphere [52]; even the nonlinear development of self-focusing has been followed for that case by a numerical simulation [53]. The case of self-focusing of a reflected heating wave was treated in [54]-[56]. The inhomogeneity of the ionosphere plays a very important role in all these theories.

Self-focusing is essentially a modulational instability. It can be considered as stimulated scattering of the pump wave by the irregularities into electromagnetic waves; the feedback is thermal rather than ponderomotive in nature. It will be recalled that the instability discussed in the last section could be regarded as stimulated scattering by the irregularities into Langmuir waves rather than electromagnetic waves.

The self-focusing of an obliquely incident pump wave was considered [57])and it was pointed out that powerful short wave stations almost certainly excite the self-focusing instability and therefore must be a source of ionospheric irregularities.

## 6. CONCLUSIONS

An attempt was made to describe the different ways in which ionospheric irregularities can be produced by powerful radio transmissions. No attempt was made to describe all the effects produced by such radio transmissions. One important omitted topic was the generation of VLF, ELF and ULF waves by modulating the auroral electrojet currents. Another omitted topic was that of ionospheric cross-modulation. The most important irregularities from the point of view of this Symposium are probably the short scale field-aligned irregularities discussed in section 4 and the much longer scale irregularities produced by self-focusing and discussed in section 5.

## ACKNOWLEDGEMENTS

This research was funded in part by Air Force Contract F19628-86-K-0031 to the University of California San Diego.

## REFERENCES

1. K. Nishikawa, Parametric excitation of coupled waves. 1. General formulation, J. Phys. Soc. Japan, 24 (1968) 916-922.

2. K. Nishikawa, Parametric excitation of coupled waves. 2. Parametric plasma photon interaction, J. Phys. Soc. Japan, 24 (1968) 1152-1158.

3. P. Kaw and J. M. Dawson, Laser-induced heating of plasma, Phys. Fluids, 12 (1969) 2586.

4. F. W. Perkins and P. K. Kaw, On the role of plasma instabilities in ionospheric heating by radio waves, J. Geophys. Res., 76 (1971) 282-284.

5. I. J. Kantor, High frequency induced enhancements of the incoherent scatter spectrum at Arecibo 2., J. Geophys. Res., 79 (1974) 199-208.

6. L. M. Duncan, The experimental saturation spectrum of parametric instabilities, Ph. D. Thesis, Rice University, Houston, Texas (1972).

7. P. Stubbe, H. Kopka, M. T. Rietveld, A. Frey, P. Hoeg, H. Kohl, E. Nielsen, G. Rose, C. LaHoz, R. Barr, H. Derblom, A. Hedberg, T. Hansen and O. Holt, Ionospheric modification experiments with the Tromso heating facility, J. Atmos. Terr. Phys., 47 (1985) 1151-1163.

8. J. A. Fejer, C. A. Gonzales, H. M. Ierkic, M. P. Sulzer, C. A. Tepley, L. M. Duncan, F. T. Djuth, S. Ganguly and W. E. Gordon, Ionospheric modification experiments with the Arecibo heating facility, J. Atmos. Terr. Phys., 47 (1985) 1165-1179.

9. J. A. Fejer and E. Leer, Excitation of parametric instabilities by radio waves in the ionosphere, Radio Sci., 7 (1972) 481-491.

10. J. A. Fejer, Physical processes of ionospheric heating experiments, Advances in Space Research, 6, #1, Symposium on active experiments (1986), in press.

11. J. A. Fejer and Y. Y. Kuo, Structure in the nonlinear saturation spectrum of parametric instabilities, Phys. Fluids, 16 (1973) 1490-1496.

12. F. W. Perkins, C. Oberman and E. J. Valeo, Parametric instabilities and ionospheric modification, J. Geophys. Res., 79 (1974) 1478-1496.

13. J. A. Fejer and Y. Y. Kuo, The saturation spectrum of parametric instabilities, AGARD Conf. Proc., 138 (1973) 11-1 - 11-8.

14. J. A. Fejer and M. Sulzer, The HF-induced plasma line below threshold, Radio Sci., 19 (1984) 675-682.

15. W. Birkmayer, T. Hagfors and W. Kofman, Small-scale plasma- density depletions in Arecibo high-frequency modification experiments, Phys. Rev. Lett., 57 (1986) 1008.

16. B. Isham, W. Birkmayer, T. Hagfors and W. Kofman, Observations of small scale plasma density depletions in Arecibo HF heating experiments, J. Geophys. Res., 92 (1987) 4629.

17. T. Hagfors, Incoherent scatter radar observations of the plasma line with a chirped pulse system, Radio Sci., 17 (1982) 727.

18. D. B. Muldrew and R. L. Showen, Height of the HF-enhanced plasma line at Arecibo, J. Geophys. Res., 82 (1977) 4793-4804.

19. M. P. Sulzer, H. M. Ierkic and J. A. Fejer, The role of Langmuir cavitons in ionospheric modification experiments at Arecibo, J. Geophys. Res., submitted (1988).

20. D. Russell, D. F. DuBois, and H. A. Rose, Collapsing caviton turbulence in one dimension, Phys. Rev. Lett., 56 (1986) 838.

21. P. A. Fialer, Field-aligned scattering from a heated region of the ionosphere; observations at HF and VHF, Radio Sci., 9 (1974) 923-940.

22. J. Minkoff, P. Kugelman and J. Weissman, Radio frequency scattering from a heated ionospheric volume, I, VHF/UHF field-aligned and plasma line backscatter measurements, Radio Sci., 9 (1974) 941-955.

23. G. H. Barry, HF-VHF communications experiment using man-made field-aligned ionospheric scatterers, Radio Sci., 9 (1974) 1025-1032.

24. J. A. Fejer, Generation of large-scale field-aligned density irregularities in ionospheric heating experiments, AGARD Conf. Proc., 138 (1973) 13-1 - 13-6.

25. G. G. Getmantsev, L. M. Erukhimov, N. A. Mityakov, S. V. Polyakov, V. P. Uryadov and V. L. Frolov, Aspect sensitive scattering of short wave radio signals by artificial ionospheric irregularities, Radiophys. Quantum Electron., 19 (1976) 1327.

26. A. Hedoerg, H. Derblom, B. Thide, H. Kopka and P. Stubbe, Observations of HF backscatter associated with the heating experiment at Tromso, Radio Sci., 18 (1983) 840-850.

27. A. J. Coster, F. T. Djuth, R. J. Jost and W. E. Gordon, The temporal evolution of 3-m striations in the modified ionosphere, J. Geophys. Res., 90 (1985) 2807-2818.

28. F. W. Perkins, A theoretical model for short-scale field aligned plasma density striations, Radio Sci., 9 (1974) 1065-1070.

29. V. V. Vaskov and A. V. Gurevich, Nonlinear resonant instability of a plasma in the field of an ordinary electromagnetic wave, Sov. Phys. JETP 42 (1975) 91.

30. V. V. Vaskov and A. V. Gurevich, Resonance instability of small-scale plasma perturbations, Sov. Phys. JETP, 46 (1977) 487-494.

31. S. M. Grach, A. N. Karashtin, N. A. Mityakov, V. O. Rapoport and V. Yu. Trakhtengerts, Parametric interaction between electromagnetic radiation and ionospheric plasma, Radiophys. Quantum Electrn., 20 (1977) 1254-1259.

32. A. C. Das and J. A. Fejer, Resonance instability of small-scale field-aligned irregularities, J. Geophys. Res., 84 (1979) 6701-6704.

33. B. Inhester, A. C. Das and J. A. Fejer, Generation of small-scale field-aligned irregularities in ionospheric heating experiments, J. Geophys. Res., 86 (1981) 9101-9106.

34. B. Inhester, Thermal modulation of the plasma density in ionospheric heating experiments, J. Atmos. Terr. Phys., 44 (1982) 1049-1059.

35. R. L. Showen and D. M. Kim, Time variations of HF induced plasma waves, J. Geophys. Res., 83 (1978) 623-628.

36. K. N. Graham and J. A. Fejer, Anomalous radio wave absorption due to ionospheric heating effects, Radio Sci., 11 (1976) 1057-1063.

37. J. A. Fejer and H. Kopka, The effect of plasma instabilities on the ionospherically reflected wave from a high-power transmitter, J. Geophys. Res., 86 (1982) 5746.

38. F. T. Djuth, C. A. Gonzales and H. M. Ierkic, Temporal evolution of the HF-enhanced plasma line in the Arecibo F region, J. Geophys. Res., 91 (1986) 12089.

39. M. A. Biondi, D. P. Sipler and R. D. Hake, Optical (6300) detection of radio frequency heating of electrons in the F-region, J. Geophys. Res., 75 (1970) 6421-6424.

40. S. M. Grach, N. A. Mityakov and V. Yu. Trakhtengerts, Electron acceleration in the presence of parametric heating of a bounded layer of plasma, Radiophys. Quantum Electron., 27 (1984) 766-771.

41. J. A. Fejer and M. P. Sulzer, Observation of suprathermal electron fluxes during ionospheric modification experiments, J. Geophys. Res., 92 (1987) 3441-3444.

42. W. F. Utlaut, E. J. Violette and A. K. Paul, Some ionosonde observations of ionosphere

modification by very high power, high frequency ground-based transmission, J. Geophys. Res., 75 (1970) 6429-6435.

43. S. A. Bowhill, Satellite transmission studies of spread-F produced by artificial heating of the ionosphere, Radio Sci., 9 (1974) 975-986.

44. S. Basu, S. Basu, P. Stubbe, H. Kopka and J. Waaramaa, Daytime scintillations induced by high-power HF waves at Tromso, Norway, J. Geophys. Res., 92 (1987) 11149-11157.

45. C. L. Rufenach, Radio Scintillations on stellar signals during artificial ionospheric modification, J. Geophys. Res., 78 (1973) 5611-5613.

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

47. A. Frey and W. E. Gordon, HF-produced ionospheric electron density irregularities diagnosed by UHF radio star scintillations, J. Atmos. Terr. Phys., 44 (1982) 1101-1112.

48. A. Frey, P. Stubbe and H. Kopka, First experimental evidence of HF produced electron density irregularities diagnosed by UHF radio star scintillations, Geophys. Res. Lett., 11 (1984) 523.

49. L. M. Duncan and R. A. Behnke, Observations of self-focusing electromagnetic waves in the ionosphere, Phys. Rev. Lett., 41 (1978) 998.

50. C. LaHoz, Studies of the self-focusing instability during ionospheric heating experiments, Ph. D. Thesis, Cornell University, Ithaca, NY (1982).

51. D. T. Farley, C. LaHoz and B. G. Fejer, Studies of the self-focusing instability at Arecibo, J. Geophys. Res., 88 (1983) 2093-2102.

52. F. W. Perkins and M. V. Goldman, Self-focusing of radio waves in an underdense ionosphere, J. Geophys. Res., 86 (1981) 600-608.

53. P. A. Bernhardt and L. M. Duncan, The feedback-diffraction theory of ionospheric heating, J. Atmos. Terr. Phys., 44 (1982) 1061-1074.

54. F. W. Perkins and E. J. Valeo, Thermal self-focusing of electromagnetic waves in plasmas, Phys. Rev. Lett., 32 (1974) 1234-1237.

55. B. L. Cragin, J. A. Fejer and E. Leer, Generation of artificial spread-F by a collisionally coupled purely growing instability, Radio Sci., 12 (1977) 273-284.

56. A. V. Gurevich, Nonlinear Phenomena in the Ionosphere, New York, Springer Verlag, (1978) chap. 6.1.

57. N. D. Borisov, V. V. Vaskov and A. V. Gurevich, Self-focusing radio wave instability in oblique ionosphere measurements, Geomagn. Aeron., 23 (1978) 563-566.

## DISCUSSION

**K.Rawer**

*Question of clarification:* In the context of PDI is what you call "ion-acoustic wave" an acoustic wave (in the ion gas) of rather high frequency which would then not be followed by the electrons so that electric fields should simultaneously occur (this type is-unfortunately-often called "electrostatic wave") or is it lower frequency acoustic or even gravity waves?

**Author's Reply**

The ion-acoustic waves of the PDI are of the same type as those responsible for incoherent backscatter. The ions and electrons "move" together and a small electric field assures quasi-neutrality. This electric field does increase the propagation velocity of the wave and this has the effect of reducing Landau damping (fewer ions are in synchronism with the wave).

**R.Showen**

You said that the radial electron velocity can be measured by the OTSI plasma line. Have any geophysical results from Arecibo or EISCAT appeared which use the OTSI enhancement?

Are the measured electron velocities the same for upshifted and downshifted lines?

**Author's Reply**

There have not been enough observations for a detailed geophysical interpretation. The data shown have not been published so far, but earlier data of poorer quality have been published in J.A.T.P. (*Fejer et al.*, 1985). The method of measuring the ion drift velocity has been greatly improved since that time (*Sulzer*, 1986).

The electron velocities have not been measured simultaneously for the upshifted and downshifted lines; it would be very surprising if they were different.

G.Rostoker

Over what height range can you use line-of-sight electron (DTSI) and ion velocities to infer line of sight current flow?

Also, is there any concern that the heater will perturb the region in which you infer the current flow making it unrepresentative of the surrounding unperturbed environment?

**Author's Reply**

The measurement of the line-of-sight drift velocity of electrons is made for the height where the heating frequency is close to the local plasma frequency. The ion velocity is measured for a somewhat greater height range.

Varying the heating power was found to have no influence on the measured drift velocity within the accuracy of the observations.

**P.Vila**

Are the photoelectrons which arise in the Shirp experiment (Arecibo) natural, or produced by the heating transmitter?

**Author's Reply**

They are natural: The photoelectrons in the Shirp experiments are only in the daytime. Their energies lie in the 10–30 eV energy range. They are not produced by the heating transmitter, but there is evidence that their distribution function is modified by the heating transmitter (*Fejer and Subaer*, 1987).

# A REVIEW ON RADIO STUDIES OF AURORAL E-REGION IONOSPHERIC IRREGULARITIES

C. Haldoupis
Physics Department, University of Crete,
Iraklion 714 09, Crete, Greece.

## ABSTRACT

During the last few years, high latitude E-region ionospheric irregularities have been studied extensively with ground based coherent VHF and UHF radar systems. Research in this field has been influenced greatly by the work done in the equatorial E-region because both phenomena have a good deal of common ground. The basic common property is that in both these regions of the ionosphere strong horizontal currents, carried by $\vec{E}x\vec{B}$ drifting electrons, do exist and are believed to constitute the main source of energy that sustains through plasma instability mechanisms the small scale electrostatic turbulence in the media. Recent auroral observations, however, which include new echo types (i.e. type 3 and type 4 echoes), large magnetic aspect angle effects, short lived and spatially coherent localized scattering regions, have reenforced the conviction that, contrary to equatorial electrojet, there is a number of important different physical processes taking place in the auroral plasma. As a result, it is now becoming clear that equatorial E-region irregularity theories, which have been applied directly to aurora (i.e. to a medium that undergoes much more dymanic variations subject to geomagnetic substorms, field aligned currents and very intense electrojet systems) need further modification to account for these high latitude backscatter observations. In this paper, we review recent radar studies of auroral irregularities at different frequency bands in the VHF and UHF range, compare their findings and evaluate their physical significance. Emphasis is placed on Doppler spectral studies, because the Doppler spectrum signatures and properties of the echoes emerged through the years as the main diagnostic tool in the identification and study of the instability mechanisms in the plasma. An objective of the paper is also to update some remaining questions/problems that need to be further studied and understood.

## 1. INTRODUCTION

The most pronounced ionospheric backscatter phenomena occur in the magnetic equatorial and high latitude (auroral) E-regions, where it is known from magnetometer measurements that strong electrojets are present. This is because of the unique properties of ionospheric plasma at E-layer's maximum near 105 km, where ion motion is controlled by collisions with neutrals because $\omega_i \ll \nu_i$ and electrons are entirely magnetized because $\Omega_e \gg \nu_e$, which leads to strong $\vec{E}x\vec{B}$ Hall currents to set up in the presence of an appreciable ambient electric field. This picture applies in both the equatorial and auroral E-regions, although the field geometry magnitude and dynamics are different. At present, it is widely accepted that the electrojets play an active role as energy suppliers in generating and sustaining the electrostatic turbulence in both media. It is this electrostatic turbulence that causes coherent scattering of radio waves at VHF frequencies or higher.

Short scale inhomogeneities in electron density have been studied extensively, in both the equatorial and auroral plasmas. It is fair to say, however, that our present interpretation and understanding of radio wave scattering phenomena from meter scale ionospheric irregularities has been based on the first radar studies of equatorial backscatter at the Jicamarca Radio Observatory in Peru (Bowles et al., 1960; 1963) and the theoretical work of Farley (1963) as well as on subsequent important contributions originated mainly at Cornell University. Radio auroral research has benefited a lot from equatorial progress, a fact that came as a natural consequence. On the other hand, direct application of equatorial knowledge to radar aurora, which is a much more complex and dynamic phenomenon of direct magnetospheric control, constitutes in many ways simplification of the problem. Therefore, it is not surprising that there is now an increasing list of observational facts in radar aurora not explained in the framework of equatorial reasoning which is founded on the idea of strictly two dimensional turbulence.

There are several comprehensive reviews on the subject of ionospheric irregularities which examine several aspects of the radar auroral phenomenon (Fejer, 1979; Fejer and Kelley, 1980; Hanuise, 1983; Fejer, 1985; Fejer and Providakes, 1987). In the present paper we review mainly recent radar results related to E-region auroral irregularities with spatial scale sizes less than 3 meters, i.e. detectable with radar frequencies $f \geq 50$ MHz. The objective is to make an integrated survey of recent observations and compare them to existing theoretical predictions. We make an effort to assess the current situation by placing the emphasis on observations with no equatorial counterpart which at present are not adequately understood and need further work to be done both, experimental and theoretical.

## 2. SOME BASICS ON SCATTERING FROM AURORAL IRREGULARITIES

Figure 1 is to define, in a general conceptual manner, the experimental problem. In Figure 1, a transmitted signal of constant amplitude and frequency is incident upon the auroral medium which acts as "natural modulator" returning to the receiver a complex signal that is both amplitude and angle modulated. In principle, the amplitude variations reflect changes of an effective irregularity scattering cross section, whereas changes in angle are due to line of sight irregularity movements. Of cource, in treating the received echoes as time varying

bandpass signal one should be aware of important limitations due to spatial effects which may cause strong interference at the receiver. Ultimately, it has been the task of radio auroral experimentalists in the last 30 years of active research to improve the resolving power of their observing systems and device various techniques to "demodulate" the back-scattered waveform and extract the returned information.




Figure 1. The incident wave is "modulated" by the scattering medium

As mentioned, in this paper we are concerned with radio wave scattering at frequencies higher than 50 Mhz, which are well above the critical frequencies reflected from the earth's ionosphere. The scattering in this case is caused by field aligned electron density inhomogeneities generated by plasma instability mechanisms. The basic physical understanding originates in the theory of scattering from a diffused ionized medium (e.g. see Farley, 1971). In this theory, for a plane wave incident upon a scattering volume $V_s$, which is large compared to the radio wavelength but still small enough to ensure statistical uniformity, there is a scattered wave whose the electric field is

$$E_s(t) \propto \int_{V_s} \Delta n(\vec{r},t)\, \exp(-j(\vec{k}_r-\vec{k}_s)\cdot\vec{r})\, d^3\vec{r} \quad , \qquad (1)$$

where $\Delta n(\vec{r},t)$ represents the electron density fluctuations from the mean ambient value $n_o$ and $\vec{k}_r$, $\vec{k}_s$ are the wavevectors of the incident radio signal and received scattered wave.

Equation (1) shows the scattered signal to depend on the spatial spectrum of electron density fluctuations evaluated at a particular Fourier component with wavevector $\vec{k}_r-\vec{k}_s$, which for backscatter in a literal sense, i.e. for a monostatic radar with $\vec{k}_r=-\vec{k}_s$, equals two times the radio wavevector. This comes as a result of constructive interference between wavelets originating at various phase fronts and adding coherently. Therefore, as Farley (1971) points out, the radio scattering measurements involve a spatial Fourier analysis of the scattering volume with the radar being sensitive only to one specific spacing of several plane wavefronts imbedded in an apparently random medium. This picture is illustrated in Figure 2. Notice that the sketched geometry is for the more general case of a bistatic system, for which the irregularity spacing depends also on the scattering angle $\Theta$ (e.i. $\lambda_{ir} = \lambda/2\sin\Theta/2$). The angle $\Theta$ is defined as $\Theta=\arccos(\hat{k}_r\cdot\hat{k}_s)$ and is equal to 180° for a monostatic radar.

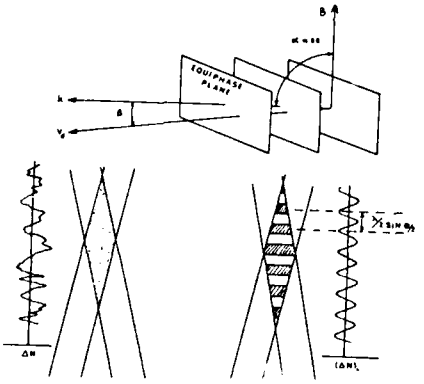



Figure 2

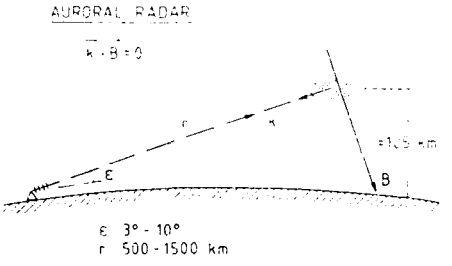



Figure 3

Figure 2. The radar is sensitive only to one spacing defined by the radio wavelength and the scattering angle (from McNamara, 1971)

Figure 3. Typical arrangement of a radar auroral experiment (Figure was provided by K. Schlegel)

The plasma irregularities are magnetic field aligned, i.e. have their phase fronts nearly parallel to $\vec{B}$. This is because in directions departing from perpendicularity, the plasma waves cannot be easily sustained due to strong Landau damping. This magnetic aspect sensitivity property of irregularities imposes certain restrictions on the experiment's location and design in general. For strong and frequent auroral echo reception, the radio system must be located well south of the scattering zone in order for themagnetic perpendicularity condition to be met. A typical monostatic (i.e. having the transmitter and receiver both at the same site) auroral radar geometry is depicted in Figure 3. In the present report, we do not review the radar experiments and techniques used by various research groups. This information can be found elsewhere in summarized form (Fejer and Kelley, 1980; Hanuise, 1983).

## 3. ELECTROJET PLASMA INSTABILITIES

Since the modern era in the study of electrojet plasma instabilities that started with the independent works of Farley (1963) and Bunemann (1963), there is a large number of papers on both linear and nonlinear theories. It is not the purpose of this report to examine all these theories and the interested reader should contact comprehensive reviews by Fejer and Kelley (1980) and Farley (1985). In this section we consider only the basic material which is pertinent to the completeness of the present paper.

In the existing linear theories of the electrojet instabilities the ambient electric field and electron density gradient are the primary driving terms of electrostatic plasma wave turbulence. Both terms can be included in the same dispersion relation, representing the two stream (Farley, 1963; Bunemann, 1963) and $\vec{E}x\vec{B}$ gradient drift (cross-field) instabilities respectively, but their significance differs depending upon the irregularity wavelength being considered. Following Rogister and D'Angelo (1970), who used a two-fluid collisional model for the electrojet plasma and assumed full magnetization for electrons and collision control for ions, one can obtain the following expressions for the oscillation frequency $\omega_k$ and the growth rate $\gamma_k$ of electrostatic plasma waves

$$\omega_k = kV_d\cos\Theta/(1+\Psi) \tag{2}$$

$$\gamma_k = \Psi/(1+\Psi)\left\{(\omega_k^2-k^2C_s^2)/\nu_i + \Omega_e\omega_k/\nu_e kL\right\} \tag{3}$$

where the term $\Psi$ is defined as

$$\Psi = \frac{\nu_e\nu_i}{\Omega_e\Omega_i}\left(1 + \frac{\Omega_e^2}{\nu_e^2}\sin^2\alpha\right). \tag{4}$$

A similar dispersion relation which in addition includes recombination effects is given by Fejer et al. (1975).

In the above equations, $\nu_e$, $\nu_i$ are the electron and ion collision frequencies with the neutrals, $\Omega_e$, $\Omega_i$ are the gyrofrequencies, k is the plasma wavenumber, $C_s=(K(T_e+T_i)/m_i)^{1/2}$ is the ion acoustic speed, $L=n_0(dn_0/dx)^{-1}$ is the electron density gradient scale length along the ambient electric field, $V_d$ is the relative electron-ion drift speed $\vec{V}_d=\vec{V}_e-\vec{V}_i$, which often is taken equal to electron Hall drift $\vec{E}x\vec{B}/B^2$. In addition to these terms, there are two directional controls built into linear theory, one represented by the "flow angle" $\Theta = \arccos(\hat{k}\cdot\hat{V}_d)$ and the other by the magnetic aspect angle $\alpha$ defined as the angle between the propagation vector $\vec{k}$ and the direction perpendicular to $\vec{B}$. It turns out that the magnetic aspect angle control is indeed quite strong because the driving terms must take unrealistically large values in order for the instability to set up and be maintained at angles $\alpha$ higher than a fraction of a degree (e.g. see Farley, 1963; Wang and Tsunoda, 1975; Schlegel and St-Maurice, 1982; Fejer et al., 1984a). This is equivalent of saying that the resulting turbulence is two dimensional with the irregularity propagation restricted in the plane perpendicular to the earth's magnetic field.

Although the fluid approximation leads to comprehensive, closed form analytical expressions, it does not apply for short wavelengths down to ion gyroradius sizes (i.e.~1 m at electrojet heights). On the other hand, kinetic theories are suitable for shorter wavelengths down to few centimeters. There are several kinetic theories developed for E-region electrojet instabilities (e.g. Farley, 1963; Lee and Kennel, 1973; Ossakow et al., 1975; Schlegel and St-Maurice, 1982; Schlegel, 1983) which arrive in comparable results with fluid theory, even at scale sizes as short as 1 m. This means that for the majority of radar observations, i.e. at frequencies less than 150 MHz, fluid theory is adequate. For shorter wavelengths, however, one has to take into consideration corrections.

In the absence of density gradients (i.e. $L\to\infty$) the equations reduce to the Farley-Bunemann instability which requires for the plasma to become unstable that $V_d\cos\Theta\geq C_s(1+\Psi)$. This mechanism can account for direct generation of short scale plasma waves inside an unstable cone where $V_d\cos\Theta\geq C_s$ and in directions close to perpendicular to $\vec{B}$, and explain some properties of type 1 echoes. These echoes have narrow Doppler spectra with mean frequency shifts corresponding to wave phase velocities at, or above, the ion acoustic speed in the plasma (e.g. see Cohen and Bowles, 1967; Balsley and Ecklund, 1972; Haldoupis and Sofko, 1976; Moorcroft and Tsunoda, 1978, among others). The role of the density gradient term on the two stream instability has been examined for the equator by Farley and Fejer (1975) and for aurora by Fejer et al. (1984a), and found to be important for longer wavelength waves, i.e. with $\lambda\geq 20$m.

When $V_d\cos\Theta < C_s(1+\Psi)$, the unified two stream and gradient drift theory fails to account even for the excitation of meter scale irregularities with phase velocities less than $C_s$. These are known to be associated with the second distinct category of equatorial and auroral echoes, named type 2, which are observed at large flow angles in directions nearly perpendicular to the electrojets. The difficulty of generating type 2 irregularities was resolved by a two step mechanism proposed by Sudan et al. (1973), in which wave energy cascades from longer wavelength primary waves to secondary short wavelength ones by the combined action of electric fields and density gradients of the primary waves. In essence, Sudan et al, (1973) applied the linearized fluid theory for indirect generation of secondary irregularities by taking simply a sinusoidal variation for the electron density gradients of long wavelength primaries and replasing $V_d$ in Equation (2) by

$$V_{ds} = -\frac{\nu_i}{\Omega_i}\frac{V_d}{1+\Psi}A\sin\varphi \tag{5}$$

where $V_{ds}$ is the relative electron-ion drift velocity caused by the polarization electric fields set up within the primary wave, A is the maximum wave amplitude (i.e. $\delta n/n_0 = A\sin\varphi$) and $\varphi$ is the phase angle in the frame of reference of the wave.

A better physical insight of Equation (5) can be obtained if we consider the local polarization fields and currents associated with a primary wave, as discussed by Sato (1973). Figure 4 illustrates these quantities within a primary sinusoidal wave (two stream and/or gradient drift) propagating along the x-direction. In region A of the wave crest the conductivity in the plasma would be somewhat higher than in neighboring region B of electron density trough. This, under the action of the primary E-field will cause alternating polarization fields to set up (e.g. Bostrom, 1973), as shown in Figure 4. In turn, these secondary fields will drive Pedersen currents along the x-direction and stronger Hall currents parallel and antiparallel to the primary destabilizing terms. All currents will be stronger in region A than B, due to higher conductivity but the normal current component at the interface between the two regions must be continuous. This requires that

$$\vec{J}_{HA} + \vec{J}'_{PA} = \vec{J}_{HB} + \vec{J}'_{PB} \quad , \tag{6}$$

where primed symbols refer to polarization field currents. By taking $n_A \simeq n_0 + \delta n$ and $n_B \simeq n_0 - \delta n$, we arrive in the following expression for the polarization fields

$$E_p = \pm E(\sigma_H/\sigma_P)(\delta n/n_0) \quad , \tag{7}$$

where $\sigma_H/\sigma_P$ is the local Hall to Pedersen conductivity ratio. Under the assumption that $\vec{E}\times\vec{B}$ electron drifts prevail in the region, the polarization drifts $V_{ds}$ which drives the secondary instabilities are expressed by

$$V_{ds} = \pm V_d(\sigma_H/\sigma_P)A\sin\varphi \quad . \tag{8}$$

The last equation reduces to Equation (6) of Sudan et al. (1973) after substituting for the conductivity ratio $\sigma_H/\sigma_P \simeq (\nu_i/\Omega_i)/(1+\psi)$.

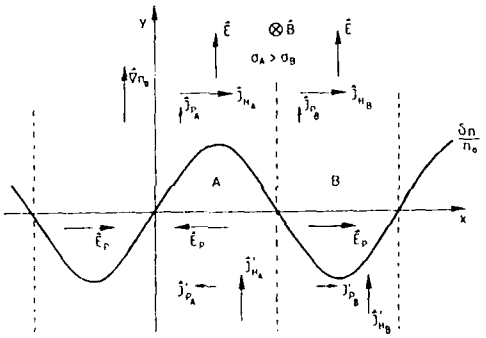



Figure 4. Secondary fields and currents within a primary electron density wave which is generated by the two stream / gradient drift instabilities.

By substituting $V_d$ by $V_{ds}$ and $L^{-1}$ by $k_p A\cos\varphi$ in equations (2) and (3), one obtains the corresponding equations for the oscillation frequency and growth rate of secondary waves

$$\omega_{ks} = k_s \frac{\nu_i}{\Omega_i} \frac{V_d\cos\vartheta}{(1+\psi)^2} A\sin\varphi \tag{9}$$

$$\gamma_{ks} = \frac{\psi}{(1+\psi)} \left\{ \frac{1}{\nu_i}(\omega_{ks}^2 - k_s^2 C_s^2) + \frac{\Omega_e}{k_s \nu_e}\omega_{ks} k_p A\cos\varphi \right\} \quad . \tag{10}$$

where subscripts p and s stand for primary and secondary, respectively. If the secondary polarization fields and density gradients are strong enough, their strength depending on the ambient destabilizing terms and the primary wavelength, secondary meter scale irregularities can be generated in the crests and troughs of the primary wave and propagate in opposite directions.

The instability conditions coming from the above equations have been exploited with success in both equatorial (Farley and Balsley, 1973; Fejer et al., 1975) and auroral (Greenwald,1974; Tsunoda, 1975; Moorcroft, 1979; Nielsen et al., 1984) backscatter studies to explain the generation and some characteristics of type 2 echoes. On the other hand, this theory fails to explain important observations such as the generation of waves with phase velocities close to zero which correspond to the strongest type 2 echoes. Obviously, the above theory is only qualitative and suggestive and cannot account for the isotropic turbulence as type 2 echo observations seem to suggest. On the other hand, numerical simulation studies

(McDonald et al., 1975; Ferch and Sudan, 1977; Keskinen et al., 1979), which retained the nonlinear terms in the equations, seem to have confirmed the two step cascade process from long to short wavelength density fluctuations, as capable in producing fairly isotropic turbulence at meter scale lengths.

Linearized theories, of course, can provide some but not all the answers. For example, they can predict the irregularity generation and what the growth rates and phase velocities are only at initial stages. The final state of the process can only be described by nonlinear mechanisms that determine the saturation of wave amplitude and the propagation characteristics of irregularities. There is a fair number of nonlinear theories developed mostly for the equator, which tried to account for observations not explained by linear theory (for relevant references see Fejer and Kelley, 1980).

Recently, there were important theoretical developments on the subject of electrojet turbulence put forward in the form of "unified nonlinear theories" by Sudan (1983), Robinson (1986) and St-Maurice (1987). In all these theories, the nonlinear interaction of electrostatic waves with the ambient plasma is recognized as an important process in understanding the physics of electrojet irregularities. In Sudan's work, a theory of strong turbulence is applied which predicts the irregularity spatial spectrum and explains several characteristics of type 2 echoes. Also, by exploiting the concept of "orbit diffusion", introduced first by Dupree 1968, Sudan developed a model which explained the stabilization of Farley-Bunemann instability spectrum and the long standing problem of constant phase velocity of type 1 echoes.This results from enhanced diffusion due to an anomalous increase of effective electron collision frequency, caused by the wave electric fields which interact with the electrons and purturb their orbit in a random manner. For a critical review of Sudan's theory, see Farley (1985).

Robinson (1986) extended the wave-enhanced diffusion ideas to electron wave heating phenomena in the auroral plasma, and found the phase speeds of Farley-Bunemann waves as a function of electron drift velocity. More recently, St-Maurice (1987) presented a unified view of the work that has been done to date on the subject of anomalous resistivity and wave heating in the unstable electrojet region, and presented several improvements. At present, researchers in the field are in the process of assessing the implications of these theoretical ideas to use them in the interpretation of observations. We shall refer to these theories again, later on.

## 4. OBSERVATIONS OF TYPE 1 AND TYPE 2 IRREGULARITIES

The terms type 1 and type 2 were originally used in equatorial E-region studies to identify two distinct irregularity types defined by the radar observing geometry and their Doppler spectrum signatures. The last few years, the same terms were also applied in auroral back-scatter studies with the same meaning as in the equator. In both cases, type 1 and 2 echoes are believed to be associated to primary and secondary plasma density waves, respectively, generated by the plasma instability mechanisms described in the previous section. In the following we review prominent auroral observations related to these irregularity types and compare them to theory.

### 4.1 Doppler spectrum signatures

Compared to equator, there is a great deal more complexity and variability in auroral Doppler spectra which reflects simply the dynamic nature of high latitude phenomena. The two prominent categories of spectra (type 1 and 2) in radar aurora,are sketched in Figure 5. The dashed line spectra indicate that they can be observed with either positive or negative Doppler shift polarity. Often as a result of temporal and spatial averaging, a fair percentage of the spectra can be a mixture of the two elementary types. Notice that type 1 are narrow with peaks at a preferred frequency shift band near the ion acoustic speed, whereas type 2 are peaked at sub-ion acoustic speeds and are much broader suggesting the existence of strong plasma turbulence. These spectral types have been reported by a number of researchers over a wide range of frequencies, i.e. in the 50 MHz band (Balsley and Ecklund, 1972, Greenwald and Ecklund, 1975; Haldoupis and Sofko, 1976; 1978a; Reed, 1980; Providakes, 1985), in the 150 MHz

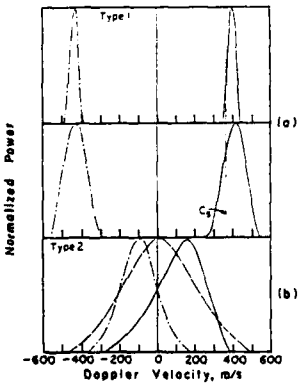



Figure 5. Drawings of typical power spectral signatures of type 1 and type 2 radio auroral echoes.

(Nielsen et al., 1984; Haldoupis et al., 1984; Schlegel et al., 1986), and in the UHF band (Abel and Newell, 1969; Tsunoda, 1976; Moorcroft and Tsunoda, 1978). Typical examples of these two spectral type observed at 50, 150 and 400 MHz are shown in the following Figure 6.

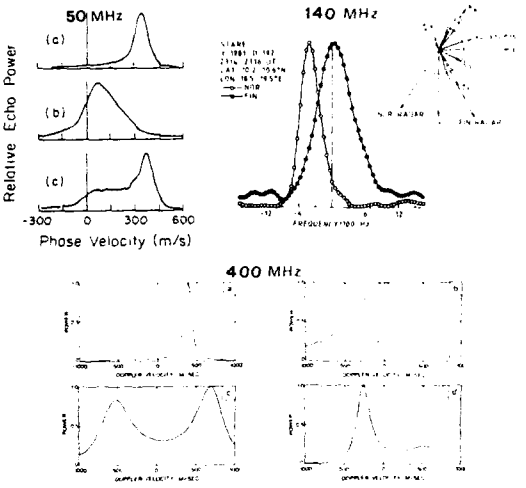



Figure 6. Typical examples of type 1 and type 2 Doppler spectra observed at different irregularity wavelengths (or radar frequencies). 50 MHz :from Reed (1980), 140 MHz: Haldoupis et al. (1984), 400 MHz : Moorcroft and Tsunoda (1978).

In general, the spectrum observations confirmed that plasma instability mechanisms which operate at the equatorial E-region are also present in the auroral electrojet as well. A direct verification was provided by the recent results of Nielsen et al. (1984) and Haldoupis et al.(1984) obtained with STARE, a twin radar system which allowed concurrent observation of the same electrojet volume at small and large flow angles. Although the spectral resolution was poor, the STARE results proved the simultaneous existence of both irregularity types (primary and secondary) in the same volume and verified the pronounced azimuthal anisotropy of Doppler spectrum properties.

Besides the spectral observations, there is also a number of other observations that favor the two stream and gradient drift instability mechanisms. These include: 1)The observation of a threshold electric field corresponding to an electron drift near the ion acoustic speed in the plasma (Tsunoda and Presnell, 1976; Moorcroft, 1979; Cahill et al., 1978). 2)The close association of backscatter properties to the electrojet strength and location (Greenwald et al., 1975; Siren et al., 1977; Haldoupis et al., 1982). 3)Some positive evidence on the destabilizing role of horizontal electron density gradients, obtained in situ from rocket measurements (Pfaff et al., 1984), ground based 50 MHz interferometric and all sky measurements (Providakes et al., 1985), and STARE spectral measurements (Haldoupis et al., 1985a). The density gradient role, however, has not yet been fully investigated.

## 4.2 Electron drift velocity and irregularity phase velocity

The radar estimates the phase velocity (or angular frequency) of plasma waves. For type 1 echoes, linear theory predicts that phase velocity follows closely variations in electron-ion drift velocity which, to a first approximation, equals the electron drift velocity $V_d = \vec{E}\times\vec{B}/B^2$ in the electrojet. This theoretical prediction and preliminary experimental evidence led several people to adopt that auroral radars could be used to deduce E-region drifts and, consequently, estimate electric fields quite accurately. This assumption seemed to work well for STARE (Greenwald et al., 1978) which has been used to estimate high latitude electric fields on a routine basis.

On the other hand, equatorial evidence on the invariability of type 1 echo phase velocity with elevation angle and its limitation to values equal the ion acoustic speed and several auroral Doppler results (e.g. Moorcroft and Tsunoda, 1978; Haldoupis and Sofko, 1978), suggested the so called cosine law relationship between phase velocity and electron drift velocity may not be valid. The first direct evidence in favor of this suggestion came from comparisons of STARE and EISCAT measurements. Nielsen and Schlegel (1985), examined simultaneous observations of wave phase velocities (STARE) and electron drift velocities (EISCAT) on a common flux tube in the E and F-regions. Their findings are summarized in Figure 7. From these results it has been confirmed that phase velocities of meter scale type 1 irregularities are limited to a value near the ion acoustic speed, in line with equatorial observations. The authors explained the variability of type 1 phase velocities as being due to alterations in ion acoustic speed caused by wave heating of electron gas.

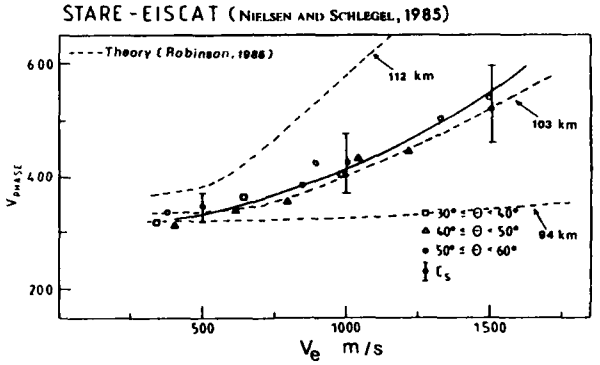



Figure 7. Experimental results (adapted from Nielsen and Schlegel, 1985) and theoretical predictions (adapted from Robinson, 1986) on the dependence of irregularity phase velocity on electron drift velocity.

A possible interpretation seems to come from Sudan's (1983) theory. In this theory, the non-linear dispersion relation results by substituting $\nu_e$ by $\nu_e + \nu_e^*$ in the linearized equations (2) and (3). $\nu_e^*$ here, represents an anomalous electron collision frequency due to momentum transfer between electrons and the plasma waves. The theory predicts that type 1 waves stabilized in this way will have phase velocities equal to the ion acoustic speed, i.e.

$$V_{ph} = \frac{\omega_k}{k} = \frac{V_d \cos\vartheta}{1+\psi+\psi^*} \approx C_s \ , \qquad (11)$$

where $\psi^* = \nu_e^* \nu_i / \Omega_e \Omega_i$. Notice that $\nu_e^*$ is anisotropic in the plane perpendicular to $\vec{B}$, depending on the wave amplitudes $\delta\vec{E}$ which maximaze along $\vec{V}_d$ and drop to zero at an angle $\vartheta_c$ defined as $\vartheta_c = \arccos(C_s(1+\psi)/V_d)$. In this way all type 1 waves inside the unstable cone would propagate at the ion acoustic speed, as it is observed.

Robinson (1986), extended Sudan's theory to include anomalous wave heating effects due to electron-plasmon (i.e. plasma wave) collisions, which would dominate over electron-neutral collisions when Farley-Bunemann waves are strongly excited. Robinson solved numerically the equations to find that $C_s$ increases monotonically with $V_d$, this dependence being stronger at heigher than lower altitudes. The altitude variations are due to changes with height of ambient plasma parameters. In Figure 7, we took the liberty to superimpose Robinson's theoretical curves on the observations of Nielsen and Schlegel (1985) for comparison purposes. Obviously, there is an excellent agreement if we assume that scattering comes, on the average, from heights near 104 to 105 km.

To get a feeling on the magnitude of $\nu_e^*$, let us consider the results shown in Figure 7 and apply Equation (11). If we take $V_d = 1000$ m/s (i.e. E = 50 mV/m), we find from the best fit curve of Nielsen and Schlegel (1985) that $C_s = 410$ m/s. Next, if we assume that scattering comes on the average from 104 km ($T_e = T_i = 212$ K and $\nu_i = 2.3 \times 10^3$ s$^{-1}$) an electron temperature $T_e = 410$ K is inferred from the above value of $C_s$ and consequently. an electron-neutral collision frequency $\nu_e = 4.6 \times 10^4$ s$^{-1}$. To compute $\nu_e$ we used the equations given by Schunk and Nagy (1978) and the CIRA 1972 model atmosphere. Finally, after substituting to Equation (11) we get $\nu_e^* = 25\nu_e = 1.2 \times 10^6$ s$^{-1}$. If we use the height of 112 km ($T_e = T_i = 260$ K, $\nu_i = 0.7 \times 10^3$ s$^{-1}$) we obtain for the same $C_s$ a value for $\nu_e = 1.1 \times 10^4$ s$^{-1}$, and finally $\nu_e^* = 280\nu_e = 0.9 \times 10^6$ s$^{-1}$.

Indirect evidence on the magnitude of $\nu_e^*$ has been provided by Primdahl and Bansen, (1985) and Primdhal (1986), who were able to explain in situ measurements of electrostatic wave potential fluctuations by using Sudan's theory and taking values for $\nu_e^*$ in the range between 0.8 to $2.5 \times 10^6$ s$^{-1}$. Recently, Igarashi and Schlegel (1987), who measured with EISCAT considerable electron temperature enhancements in auroral E region, found an excellent agreement with theoretical estimates of electron gas heating by assuming an average density perturbation of 3.5 % (heating theory of St-Maurice et al, 1981), or an anomalous electron collision frequency $\nu_e^*$ of 1.0 to 2.0 $\times 10^6$ s$^{-1}$ (theory of Robinson, 1986). The quoted values for $\nu_e^*$ are all in the range from a few percent to 25 % of the electron gyrofrequency ($\Omega_e = 10^7$ s$^{-1}$). *Can $\nu_e^*$ take such large values without violating the basic assumption that electrons are strongly magnetized?* This question has been addressed recently by St-Maurice (1987) who estimated the upper limit for $\nu_e^*$ to be less than about 10 %, by using known estimates of full turbulence limits. He concluded that, if this is indeed true, then $\nu_e^*$ may in fact be overestimated by a Sudan-type theory.

## 4.3 Echo Occurrence at Large Magnetic Aspect Angles

One of the most important discrepancies in the nature of equatorial and auroral backscatter is related to the magnetic aspect properties of irregularities. Contrary to equatorial case, radio auroral type 1 and 2 echoes do exist, although weaker, at magnetic aspect angles

much higher than one degree allowed by theory. There have been many radar auroral experiments, which have observed routinely strong radio auroral echoes coming from aspect angles in the range between 1° to more than 15° (e.g. Leadabrand et al., 1965; McDiarmid and McNamara, 1967; McNamara, 1971; Hofstee and Forsyth, 1972; Ecklund et al., 1975; Mitchel and Brown, 1976; Sofko et al., 1983; Koehler et al., 1985; Providakes, 1985; Haldoupis et al., 1986).

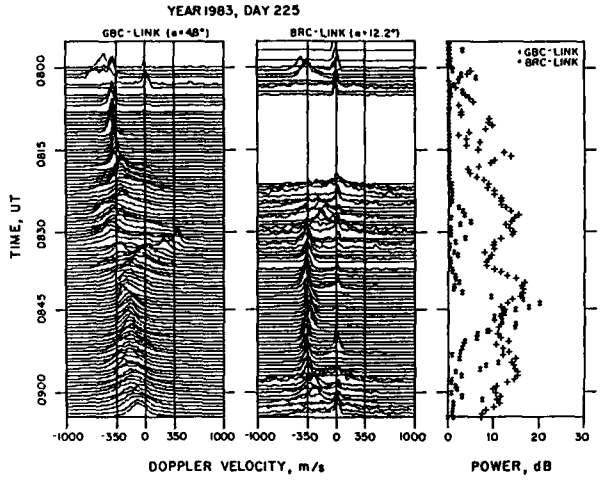



Figure 8. Simultaneous sequential Doppler spectra from two different azimuths show that type 1 and type 2 echoes are observed at quite large magnetic aspect angles ( 5° and 12°).

Recently, Haldoupis et al. (1987) analyzed a large data base compiled from several experiments carried out in Central Canada by the University of Saskatchewan. The measurements were made with high resolution CW Doppler systems at 10 different aspect angles ranging between 1° and 14°. The evidence suggested that 3 meter primary and secondary plasma waves with $k_{\parallel}/k_{\perp}$ wavenumber ratios perhaps larger than 0.25 do exist in the auroral plasma on a continuous basis. Figure 8 shows several examples of both type 1 and type 2 echoes observed simultaneously from two different azimuths at fairly large aspect angles, 5° and 12°. The aspect sensitivity statistics of type 1 and 2 echoes, presented by Haldoupis et al. (1987), are summarized in Figure 9. Notice, the basic spectral properties of 50 MHz echoes follow only minor changes with aspect angle in the 1° to 15° range. On the other hand, Ogawa et al. (1980) reported a significant decrease of type 1 phase velocity with aspect angle in the 1° to 5° range for 50 MHz backscatter. This result however, was based on limited data and it is not clear if they dealt with type 1 irregularities only.

The linearized theories of both primary and secondary electrojet instabilities fail to explain the generation of irregularities at aspect angles higher than a degree or so. A mechanism which could extend electrostatic turbulence at higher aspect angles is some kind of anomalous resistivity that introduces electron collision frequency enhancements in the dispersion relation. The destabilizing role of electron collisions is well recognized in current driven instabilities (e.g. D'Angelo, 1973). Such a mechanism was proposed by Volocevich and Liperovskiy (1975). Haldoupis et al. (1986) have shown that one needs very large collision frequencies of the order 5 to 15 % of

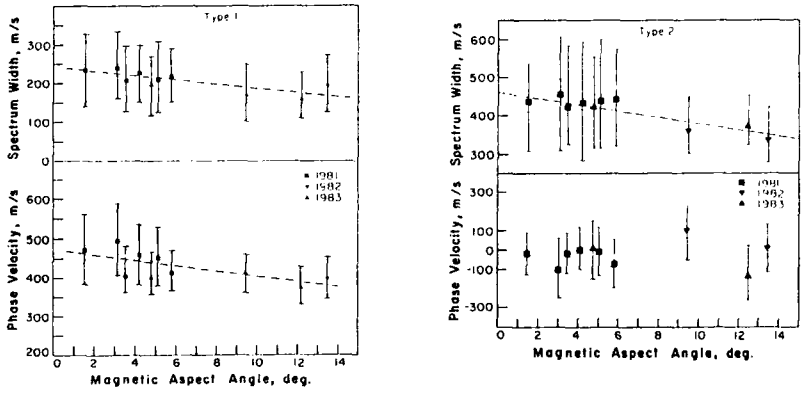



Figure 9. Type 1 and type 2 spectral statistics for different magnetic aspect angles at 50 MHz.

the electron gyrofrequency, to explain scattering at large aspect angles. There seems to be no theory to justify such large electron collision frequencies. A Sudan-type theory, discussed in the previous section, is applicable only for $k_{\parallel} \simeq 0$. Therefore, the concept of anomalous $\nu_e$ enhancements may be suitable for explaining effects at small aspect angles, as it has been demonstrated recently by Nielsen (1987) for 140 MHz data, but is inadequate for backscatter at aspect angles larger than a couple of degrees. At present, we have not yet a convincing theoretical interpretation of auroral backscatter properties at large magnetic aspect angles.

4.4 Irregularity Scattering Cross Sections

In the radar equation (e.g. see Moorcroft, 1987a), the received signal strength is directly proportional to the scattering cross section (SCS) of the irregularities. The latter is a measure of absolute mean level of electron density fluctuations along the radar line of sight and only at the specific spacing which equals half the radar wavelength, i.e.

$$\sigma_k \propto \langle |\Delta n(2k_r)|^2 \rangle \qquad (12)$$

where $k_r$ is the radar vector wavenumber and $\Delta n(2k_r)^2$ is the mean spatial Fourier component, at the specific wavelength $\pi/k_r$, of electron density fluctuations in the scattering volume. In principle, one can measure the absolute radar scattering cross section and derive information about the rms level of fluctuations in the medium. In practice, however, the difficulty in determining the various quantities entering the radar equation makes measuring $\sigma_k$ a very tricky problem. On the other hand, the relative scattering cross section can easily be estimated if one adopts some reasonable normalization.

Accumulated experimental evidence shows that $\sigma_k$ depends upon a number of parameters such as electron drift $V_d$, ambient electron density, flow angle $\vartheta$ and magnetic aspect angle $\alpha$. In other words, $\sigma_k$ becomes a complicared function, $\sigma_k = f(V_d, n_0, \vartheta, \alpha)$, whose form is not known. At present, the existing observations provide only simple empirical relationships, whereas a unified theoretical treatment of the subject is far from complete.

Magnetic aspect angle. The dependence of $\sigma_k$ on $\alpha$ is a well known phenomenon since the beginning of radio auroral research. The decrease of backscatter intensity with magnetic aspect angle, known as aspect sensitivity (AS), is expressed in dB/degree. The overall picture on AS remains unclear and somewhat contradictory, because various experiments yielded widely different results. From the existing publications, one may in general conclude that echoes at 50 MHz are less aspect sensitive than those at 150 or 400 MHz. At 50 MHz the echo intensity drops with an average rate of 1 to 3 db/deg while at higher frequencies the rate is near 10 dB/deg (e.g. Leadabrand, 1965; Chesnut, 1968; McDiarmid, 1972; Ecklund et al., 1975; Koehler et al., 1985; Haldoupis et al., 1986). Apparently, this difference emmanates from the fact that longer scale electrostatic turbulence is more easily generated compared to shorter scale one. In order to explain the aspect sensitivity, Moorcroft (1985) introduced a scattering model consisting of an assembly of wavelike irregularities, each with a gaussoidal envelope, that added up having an effective scattering coefficient as a function of aspect angle. Although Moorcroft had taken into consideration all possible contributions to aspect sensitivity (e.g. magnetic field line distortion and ionospheric refraction), he concluded, the experimental observations required scattering models consisting of irregularities elongated along the magnetic field by only a few plasma wavelengths which is physically unreasonable.

Flow angle and drift velocity effects. Naturally, the well known azimuthal Doppler spectrum anisotropy, caused by the presence of two distinct irregularity types that dominate entirely different sectors in the observing plane relative to the destabilizing drift, is expected to apply on irregularity SCS as well. The first systematic study of auroral SCS in the azimuthal plane was carried out by Andre (1983) at 140 MHz with the STARE dual radars. Andre's results confirmed an anisotropy in the signal strength with flow angle, with a density fluctuation minimum at $\vartheta \simeq 90°$ and an increase towards $\vartheta \simeq 0°$ over the drift velocity range up to 1000 m/s.

Haldoupis and Nielsen (1984) studied the relative SCS angular dependence of 140 MHz backscatter using Doppler spectra to identify the irregularity types. In their results, type 1 primary irregularities, have highly anisotropic SCS as compared to secondary 1 m waves. The strength of primaries is strongly dependent on flow angle with the intensity decreasing from 0.3 to 0.6 dB/deg as $\vartheta$ increases; on the other hand, the strength of secondaries is rather weak and insensitive to angular variations, in line with the concept of isotropic turbulence. These findings agree with similar results reported by Ierkic et al. (1981) for 50 MHz equatorial backscatter. Recently, Matin and Jones (1987) reported a detailed study on flow angle dependence of 1 meter irregularities based on dual radar SABRE measurements, which show that the problem is more complex. They found that, although for small drift velocities there is isotropic behaviour, at higher drifts there are anisotropies developing in the observing plane which are not symmetric showing unexpected subsidiary minima at directions along the drift, i.e. at directions where one would have expected the signal to maximize.

As pointed out by Fejer and Providakes (1977), Sudan's (1983) theory predicts for type 1 waves at electrojet altitudes with $k_{\parallel} \simeq 0$, that

$$\sigma_k \propto \left\{ (V_d \cos\vartheta - C_s) / C_s \right\}^2 \qquad (13)$$

Notice, that according to the theory $C_s$ is also depending on $V_d \cos\vartheta$ due to anomalous electron gas heating as it has been discussed in section 4.2. Equation (13) implies that wave saturation amplitudes follow a nonlinear increase with radial drifts $V_r = V_d \cos\vartheta$, in general agreement with the observed flow angle anisotropy. The only semi-quantitative experimental study on the dependence of relative SCS on $V_d \cos\vartheta$ has been reported by Haldoupis

et al. (1984) for 1 meter waves with $k_{\parallel} \simeq 0$, using STARE spectral data. In the analysis, they used a large data base from different events and assumed the effect on $\sigma_k$ from other parameters to average out to a large extend. In addition, they have dealt with electric fields less than 40 mV/m (i.e. electron drifts less than 800 m/s) which are believed to be estimated accurately with STARE (Nielsen and Schlegel, 1983; 1985). From Figure 10a , showing the results of Haldoupis et al. (1984), it is seen that on the average $\sigma \propto (V_d \cos\theta)^{2.5}$. To compare with theory, we plot in the same figure the result of Equation (13) using some appropriate arbitrary normalization. The comparison shows rather poor agreement for radial velocities below 500 m/s i.e. in the range where most of type 1 echoes are observed.

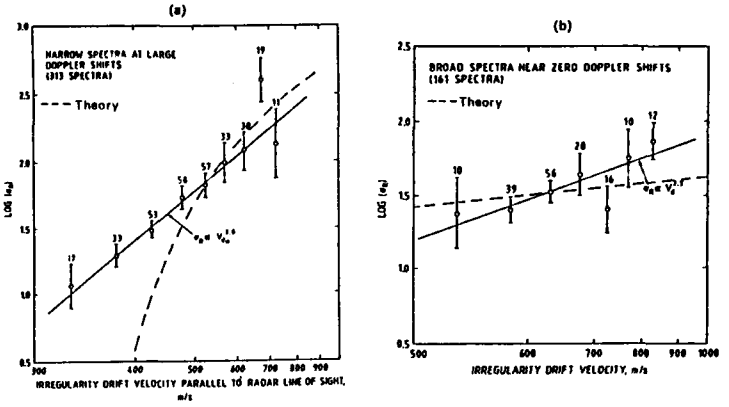



Figure 10. Relative scattering cross section dependence on drift velocity for type 1 (Fig. 10a) and type 2 (Fig. 10b) 140 MHz echoes. The dashed lines were based on theoretical predictions (Sudan, 1983) using arbitrary normalizations.

For secondary irregularities, the turbulence theory of Sudan (198 ) shows the saturation amplitudes to depend on both ambient destabilizing terms $V_d$ and L in the following way, as shown by Farley (1985),

$$\sigma_k \propto L^{-4/3} V_d^{2/3} \quad . \tag{14}$$

The STARE data study, seen in Figure 10b, show on the average that $\sigma \propto V_d^{2.3}$ , a result that is in good agreement with a similar dependence found at 50 MHz for the equator type 2 echoes (Balsley, 1969; Farley and Balsley, 1973) and with numerical simulation results (Keskinen et al., 1979). The agreement between theory and the STARE observations is not as satisfactory, as shown in Figure 10b where we superimposed the result of equation (14) assuming $L \simeq$ const. For a detailed discussion on this point see Farley (1985). As mentioned $\sigma_k$ depends on a number of parameters, therefore the above comparisons are only indicative and should not be taken as conclusive.

Ambient electron density. We have just seen that a strong dependence of $\sigma_k$ on $V_d$, or the ambient E field, does exist. On the other hand, there is a number of studies both at 50 MHz (Greenwald et al., 1975; Siren et al., 1977) and 140 MHz (Haldoupis et al., 1982; Uspenskiy et al., 1983; Starkov et al., 1983) suggesting that $\sigma \propto n_0^2$. Apparently, The backscatter strength under some sircumstances is governed by the ambient E field and under some sircumstances by the ambient electron density. The $n_0$ control seem to be dominant at drift velocities well above the required threshold for type 1 waves. The reason for $\sigma_k$ to be at times proportional to $n_0^2$ and the competing roles of $n_0$ and E field, are not clear. It is interesting to point out that the existing theories do not include the ambient electron density in the calculations of wave saturation amplitudes. Certainly, the role of $n_0$ in the backscatter process is not well understood.

Absolute scattering cross sections. Oksman et al. (1986), were the first to attempt some specific calculations of absolute SCS of radar aurora, utilizing data from several radars, to calculate mean fractional electron density fluctuation amplitudes. Recently, Moorcroft (1987a) conducted a thorough study on absolute SCS by exploiting pubished material on callibrated backscatter measurements to estimate absolute scattering coefficients per unit volume in over a range of frequencies from 30 to 1210 MHz. In his calculations, Moorcroft used a 10 km width for the scattering layer and assumed no magnetic aspect angle effects. The estimated SCS coefficients are illustrated in Figure 11. Notice in Figure 11a, the scattering coeficients vary over 3 orders of magnitude at any frequency, and a definite trend for cross sections to decrease with frequency.

By using a number of arguments based on experimental evidence, Moorcroft (1987a) concluded that the strongest events in each distribution of Figure 11a are due to primary (type 1) irregularities. He produced Figure 11b which shows estimates of maximum SCS as function of radar frequency or irregularity wavelength. The straight line is a least squares fit to all points except for 30 and 1210 MHz data which were excluded for a number of reasons discussed in detailed by Moorcroft (1987a). In the range between 50 and 800 MHz, i.e. $\lambda_{\perp}$ between 3 m and 0.2 m, the frequency dependence for waves propagating along the destabilizing flow and perpendicular to $\vec{B}$ can be described well by an absolute scattering cross section per unit volume

that follows the expression

$$\sigma_v = 3\times10^{-7}k^{-2.25} \quad m^{-1} \; . \qquad (15)$$

The dotted lines in Figure 11b indicate a range uncertainty of 0.35 for the power of 2.25 in the above equation. It is interesting to point out that these results are in fair agreement with Sudan's (1983) nonlinear theory which suggests a $k^{-2}$ dependence for the saturation amplitudes of type 1 short wavelength irregularities.

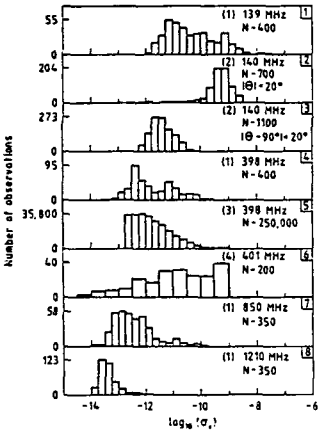



Fig. 11a

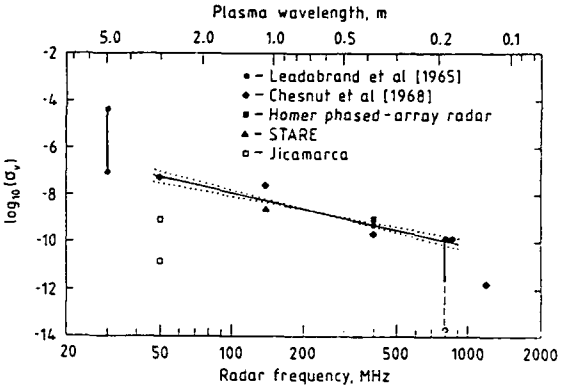



Fig. 11b

Figure 11. Both figures are from a recent paper by Moorcroft (1987a).(a): Distribution of scattering coefficients at various frequencies. (b): Estimates of maximum absolute scattering coefficients per unit volume as function of radar frequency.

## C. ALTITUDE CHARACTERISTICS OF RADIO AURORA

First, we should point out that most radar experiments could not provide adequate height information, to be used in the interpretation of data, because of broad antenna beamwidth in elevation. This was an important disadvantage *because there are several reasons to believe* that height effects do exist and play a crucial role in the physics of the phenomena. Ambient plasma parameters (e.g. temperatures, collision frequencies, electron density and density gradients) that affect the instability mechanisms and the magnitude and direction of electrojet, do change appreciably with altitude. For example, Figure 12 shows important height differences in the theoretical dependence of threshold electric field for two stream waves at various frequencies. Furthermore, there is enough evidence showing sizable dynamic alterations of plasma parameters *to take place within a few kms altitude* in the electrojet region. We show such an example in Figure 13 reproduced from a paper by Igarashi and Schlegel (1987). This anomalous electron heating in the auroral electrojet, discovered first by Schlegel and St-Maurice (1981), is attributed to plasma particle interactions with strong Farley-Bunemann waves and believed to act as a feedback element in the saturation of two stream instability.

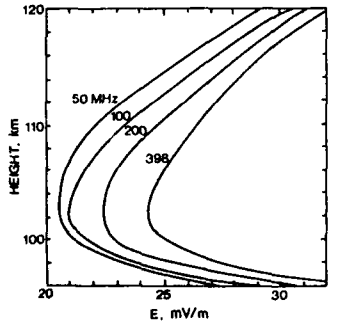



Figure 12 (from Moorcroft, 1979)

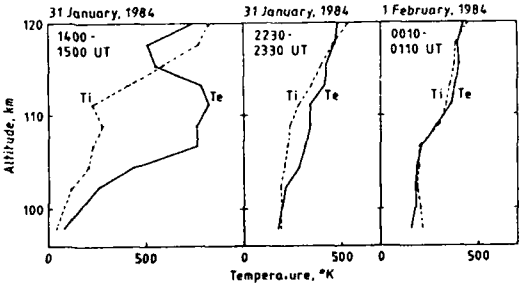



Figure 13 (from Igarashi and Schlegel, 1987)

The auroral echoes are believed to come from an E region layer centered near 107 km and having a variable thickness from less than 2.5 km to possibly more than 20 km. These numbers came from radar backscatter altitude estimates (e.g. see Umwin, 1972). Also recket measurements detected, with ac electron density and E field probes, a layer of strong electrostatic turbulence located between 95 and 120 km (e.g. see Kelley and Mozer, 1973; Bahnsen et al., 1978; Ogawa et al., 1976; Pfaff et al., 1984). The first detailed measurements of radar

auroral altitude effects, reported by Unwin and Johnston (1981) with a sea interferometer radar in south New Zealand, revealed considerable altitude structuring within the electrojet layer. At times, the observed two scattering layers, at about 106 and 112 km, producing echoes with different spectral characteristics. A second detailed study using the same sea interferometer technique, has been conducted by Timofeev and Miroshnikov (1982) with a 90 MHz radar in Karelia, USSR. Their findings include, 1) more altitude variability and layering in the westward than eastward electrojet, 2) a pronounced east-west altitude asymmetry in both electrojet time sectors and, 3) short lived and localized double scattering layers with a thickness of about 2 km and an altitude separation of less than 10 km. They have attributed these effects on a mechanism of collective interaction of energetic auroral electron fluxes with the E-region plasma. In general, we do not know at present in what way this altitude structuring affect the scattering mechanisms of radar aurora.

Important information on height effects for 398 MHz backscatter have been obtained by Ruohoniemi and Moorcroft (1985), who used the narrow beam elevation scanning capability of the Homer radar in Alaska. In their technique, they exploited the different echo intensities from successive elevation positions at a given range to calculate the elevation angle, and therefore the altitude, of the scattering layer. They found considerable differences in the altitudes for premidnight and postmidnight backscatter periods, i.e. in the eastward and westward jets, respectively. For the east jet the echo altitudes were confined between 97 and 117 km whereas for the postmidnight period (westjet) the scatter layer was much narrower centered at lower altitudes near 103 km. The lower heights for the westward electrojet could be explained, considering the role of electron density in the scattering cross section as described by Uspensky (1985), and that the west jet is ussually centered at altitudes considerably lower than the eastward jet (e.g. see Kamide and Brekke, 1977). Further work by Moorcroft and Ruohoniemi (1987) on Doppler velocity altitude variations at 398 MHz, revealed, among other things, that low velocity echoes, presumably due to secondary waves, are more restricted in height range than echoes with ion acoustic velocities which are believed to come from primary, Farley-Bunemann waves. This interesting result of Moorcroft and Ruohoniemi (1987), which is illustrated in Figure 14, is difficult to be explained and understood without supportive observational evidence.

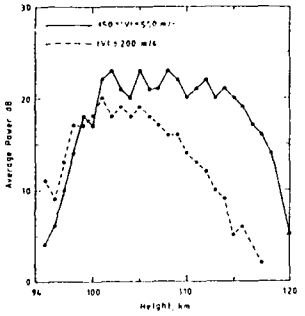



Figure 14. Average echo power versus altitude for type 1 and type 2,398 MHz backscatter. (from Moorcroft and Ruohoniemi, 1987).

At present, the existing theories are not adequate in explaining altitude observations of the type shown in Figure 14. As Moorcroft and Ruohoniemi (1987) pointed out, it is now clear that height information is essential for the physical understanding of auroral backscatter observations. Furthermore, detailed height measurements might be useful in testing existing theoretical predictions on plasma wave propagation within the unstable layer, both upward and downward along the magnetic field lines (Moorcroft, 1984; St-Maurice, 1985; Moorcroft, 1987b). For example, the waves grow as they propagate and if growth rates are larger compared to propagation times, then the strongest echoes are expected to come from the top and bottom of the unstable layer.

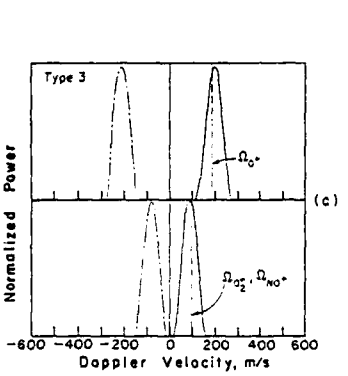
## 6. NEW ECHO TYPES AT 50 MHz AURORAL BACKSCATTER

In the previous sections we reviewed type 1 and type 2 irregularity characteristics detectable in the entire VHF and UHF range. These were interpreted in the framework of two stream, gradient drift and secondary instabilities. In this section we examine in brief two newly discovered categories of auroral radar echoes observed only at 50 MHz, i.e. with 3 m irregularities. The terms used to identify these echoes are "type 3" and "type 4", in contrast to the familiar type 1 and type 2.

### 6.1 Type 3 echoes and ion cyclotron instabilities.

Type 3 echoes (Reed, 1980; Fejer et al., 1984b; Providakes et al., 1985; Haldoupis et al., 1985b; Prikryl et al., 1987) are identified by their Doppler spectrum signature sketched in Figure 15. On the average, type 3 echoes are related to strong intensity signals having very narrow spectra peaked below ion acoustic Doppler speeds and preferentially either in the 50 to 70 Hz shift band (150 to 220 m/s) or less often in the 20 to 40 Hz shift range (60 to 120 m/s). These frequency bands are centered on values corresponding approximately to the cyclotron frequencies, or their harmonics, of the main ionic species in the plasma (i.e. $O_2^+$, $NO^+$, $O^+$). Figure 16, is an example of type 3 spectra centered near 30 and 60 Hz, observed simultaneously from different azimuths and at aspect angles 5°and 12°, respectively. It is

interesting to note that all the reported observations were made with 50 MHz radars whose beams were several degrees from perpendicularity with the magnetic field, i.e. type 3 echoes seem to require less favorable aspect sensitivity conditions than type 1 and 2 echoes. According to Haldoupis et al. (1987), type 3 echoes were seen frequently in the entire aspect angle range between 1°and 14°, i.e. with $k_{\parallel}/k_{\perp}$ ratios may be larger than 0.25. Also, type 3 echoes are observed during more disturbed rather than more quiet geomagnetic conditions.



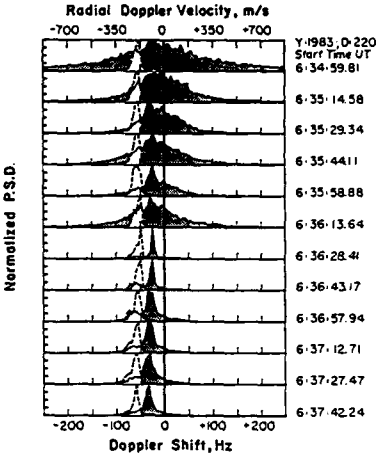




Figure 15. Drawings of typical type 3 spectra observed at 50 MHz.

Figure 16. Sequential type 3 spectral peaks at 30 and 60 Hz, observed simultaneously from different azimuths and magnetic aspect angles (from Haldoupis et al. 1985b)

These characteristics, and in particular the fixed Doppler shift near the ionic gyrofrequencies, led Fejer et al. (1984b) to propose that these echoes are due to electrostatic ion cyclotron plasma waves generated by strong field aligned currents (FAC) in the upper E region (i.e. well above the electrojet). Besides the Doppler spectrum evidence, there is some indirect evidence based on concurrent optical data, reported by Providakes et al. (1985) and by Prikryl et al. (1987), in favor of strong FAC involvement in the generation of type 3 echoes. Also, there are measurements made in situ with rocket probes which detected ion cyclotron electrostatic oscillations in the upper E region (e.g. Ogawa et al., 1981; Bering, 1983). The existing evidence, however, should be considered as necessary but not sufficient to prove the electrostatic ion cyclotron source mechanism for auroral backscatter, and further experimental evidence, especially altitude information, is needed.

The first to discuss theoretically the importance of FACs in the generation of electrostatic ion cyclotron waves were Kindel and Kennel(1971). Their theory applied to the weakly collisional upper E region and lower F region ionospheric plasma (i.e. for $\nu_i \ll \Omega_i$ and $\nu_e < k_{\parallel} V_e$, where $V_e$ is the electron thermal speed). According to the kinetic theory of Kindel and Kennel, once the threshold drift velocity is exceeded, short wavelenght waves near the ion cyclotron frequency (i. e. $\omega_r \sim \Omega_i$) become unstable and propagate outward from a current beam in directions such that $k_{\perp} \gg k_{\parallel}$. This theory however shows that ion neutral collisions have a strong stabilizing effect (even when $\nu_i/\Omega_i \simeq 0.02$), thus extreme FACs would be needed to sustain EIC waves below ~160 km.

On the other hand, D'Angelo (1973) and Chaturvedi (1976), who used fluid theory, considered the destabilizing role of electron neutral collisions and argued that EIC waves could be generated even at 120 km, where $\nu_i$ and $\Omega_i$ become comparable. Fejer et al. (1984a) generalized the fluid theory of electrojet instabilities to include ion magnetization effects and FAC generation of EIC waves at higher altitudes. Fluid theory, however predicts direct generation of EIC waves with $\lambda_{\perp} \geq 15$ m to 20 m and frequencies $\omega_r = (\Omega_i^2 + k^2 C_s^2)$, i.e. phase velocities well above the ion acoustic velocity $C_s$. Obviously, this theory, neither explains the generation of 3 m waves nor predicts the observed Doppler shifts so near the ionic gyrofrequencies. At this point, it is interesting to mention that very recently Villain et al. (1987) using the APL HF radar facility in Labrador, which scans in frequency from 8 to 20 MHz, reported observations of E region irregularities with $\lambda_{\perp}$ near 15 m which can be interpreted in terms of EIC waves produced by $NO^+$ ions, in good agreement with the above dispersion relation of fluid theory.

A kinetic study of collisional EIC waves ($\nu_e > k_{\parallel} V_e$) in the upper E region, which took into consideration the destabilizing effect of electron neutral collisions, was reported by Providakes et al. (1985). As seen in Figure 17, reproduced from the paper of Providakes et al. (1985), the theory predicts that 3-m EIC waves can have angular frequencies near the ion gyrofrequencies but for direct excitation unreasonably high electron drift velocities along the field lines are required, even higher than the electron thermal speed (~130 km/s). The required drifts at instability threshold increase immensly at larger aspect angles and, as the authors state, even the most easily excited EIC waves, occurring at 0.5 aspect angle angle, cannot be generated.

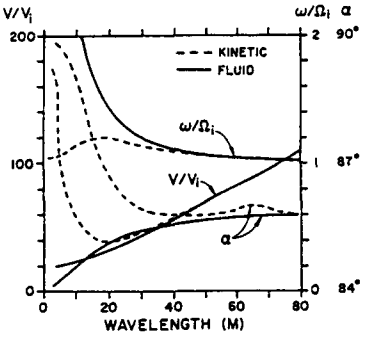



Figure 17. Theoretical results for electrostatic ion cyclotron waves at instability threshold in the upper E-region ( 150 km). $V/V_i$ is the ratio of field aligned electron drift to the ion thermal speed. (from Providakes et al., 1985).

In summary, collisional fluid or kinetic theories of FACs ion cyclotron instabilities cannot provide adequate explanations for the type 3 observations at 50 MHz. Yet the frequent observation of these waves at large aspect angles shows the theories must be modified so as to relax the threshold conditions for instability in both small and large aspect angles and to allow for phase velocities near the ion gyrofrequencies. Also, the theories should explain why type 3 echoes, contrary to type 1 and 3, are not observed by radars at frequencies higher than 50 MHz. Also careful experimental studies are needed to determine the altitude at which these echoes originate. Based on the observation that type 3 phase velocities remain fairly stable, as in the case of type 1 echoes, one could argue that type 3 echoes may come from within or near the top of the elctrojet because if they were to come from the upper E-region (i.e. say 150 km), then their Doppler shift would have been subjest to ExB plasma motions. On the other hand, if type 3 echoes come even from upper electrojet layers, can we define electrostatic ion cyclotron waves in regions where the ions are hardly magnetized?

6.2. Type 4 echoes.

A few years ago, Haldoupis and Sofko (1979) reported a new spectral signature observed with a CW bistatic Doppler system at 42 MHz, during exceptionally disturbed conditions in the postmidnight sector. An example of their observations is shown in Figure 18. The spectrum consisted of two components, a broad one centered near 500 m/s and a narrower one near 1.1 km/s. The overall duration of such events was fairly short, from tens of seconds to less than 3 to 4 minutes. The evidence suggested that short lived, lifetimes less than 50 ms, fast moving irregularities must have originated from very localized scattering regions during unusual plasma conditions.

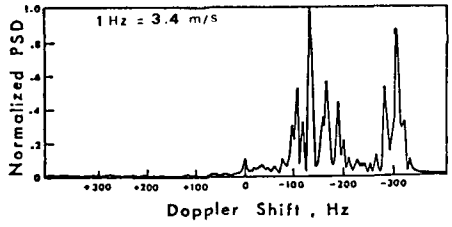



Figure 18. A typical example of the type 4 50 MHz spectral signature. This spectrum corresponds to a short data sample of 0.2 s (from Haldoupis and Sofko, 1979).

This observation has been identified as a distinctly different echo type of 50 MHz radar aurora by Providakes (1985) and Fejer et al. (1986) who observed several type 4 events with the high resolution Cornell University Portable Interferometer Radar System (CUPRI), during disturbed conditions in the posmidnight sector. On these occasions, short lived narrow spectral peaks with very large phase velocities nearing 1 km/s, are observed to dominate over a broad spectral part of type two-like but with considerably higher average Doppler velocities and spectral widths. The echo strength is often high and variable in range and time and, as shown in Figure 19, the Doppler spectrum follows dynamic changes. They also found the type 4 component to be more aspect sensitive than the broad type two-like component.

In their interpretation, Fejer et al. (1986) relied on the shape of the peaks and their time evolution to suggest that type 4 events represent a true two stream instability case, implying that at these times the ion acoustic speed is extremely enchanced due to anomalous electron heating conditions. The rare occurrence, short lifetimes, and highly variable nature of type 4 peaks, were explained as a consequence of an extremely narrow cone of primary two stream irregularities, so type 1 waves are confined along the destabilizing drift and observed only when $\cos(k V_d) \simeq \pm 1$. This interpretation implies very large electron temperatures in the range between 2000 °K and may be more than 3000 °K. As we have mentioned, E region electron temperature enhancements are known, from ionospheric scatter measurements, to occur during periods

periods of large electric fields, but very rarely we have such large values being reported in the literature.

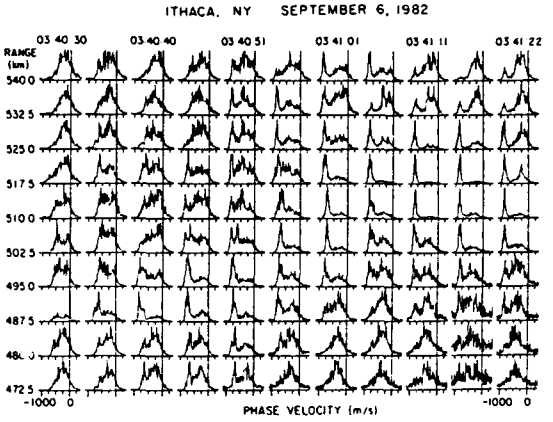



Figure 19. Temporal and spatial dynamics of type 4 spectra observed by the CUPRI radar at 50 MHz. (from Fejer et al. 1986).

We believe, the proposed interpretation for type 4 echoes is feasible but it needs more experimental testing. One may find it difficult to conceive that type 4 echoes are only due to a narrow, highly directional, cone of primary irregularities, when we know from STARE that the primary cone is always very broad (for example the Norway STARE radar observes almost always narrow type 1 spectra), even with large electron drifts. Also from the published type 4 spectra, e.g. those of Haldoupis and Sofko (1979) and Fejer et al. (1986), there seems to be a prefered frequency band close to 300 Hz for type 4 spectral peaks, something that may be not compatible with the proposed interpretation.

## 4. INTERFEROMETRY STUDIES OF LOCALIZED SCATTERING REGIONS

In the past, there have been a number of observations suggesting the existence of localized strong scatter regions in auroral plasma (Sofko and Kavadas, 1969; 1971; Unwin, 1972; Haldoupis and Sofko, 1978b; Sinclair and Forsyth, 1980). The notion of backscatter being dominated by echoes from a dynamic point-like source imbeded in an apparently random medium is intriguing, hinting that under some appropriate conditions occuring at limited volumes of space, the plasma is succeptible to instability processes that need to be studied and understood.

A new technique which can provide information on the spatial and temporal structure of localized echoing regions has been developed at Cornell University and applied succesfully the last few years to equatorial and auroral electrojets (Ierkic, 1980; Farley et al., 1981; Kudeki et al., 1981; Providakes et al., 1983; 1985; Riggin et al., 1986; Fejer et al., 1986). The first auroral interferometer results, reported by Providakes et al. (1983) using a monostatic 50 MHz coherent radar with antenna separation of 50 m, verified the existence of localized scatter sources in the auroral plasma with dimensions of less than few kilometers. Recently, Haldoupis et al. (1988) described the conversion of a bistatic 50 MHz Doppler system to an auroral interferometer designed to investigate short term dynamics of highly localized regions of backscatter with transverse sizes in the range from 0.5 to 5 km.

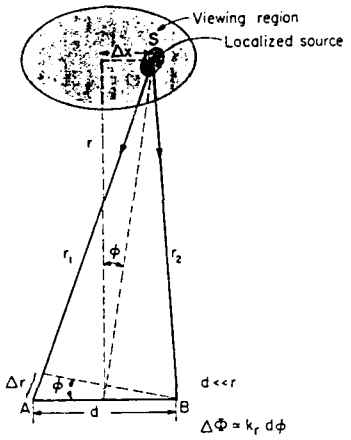
Figure 20 illustrates the basic geometry of horizontal (i.e. azimuthal) interferometry for observing a localized source within a large viewing region. If we assume the received signals result from the superposition of many wavelets originating from a random assembly of source elements having their phases $\Phi_{\omega_i}$ normally distributed, it can be shown that the complex cross spectrum is

$$S_{AB}(\omega) = \exp(jk_r d\,\bar{\Phi}_\omega)\,\exp(-k_r^2\, d^2\sigma_\omega^2/2)\ , \qquad (16)$$

where $k_r$ is the radar wavelength, d is the antenna separation distance, and $\sigma_\omega$, $\bar{\Phi}_\omega$ are the angular spread and the mean azimuthal position of the source which is identified in the Doppler spectrum by its mean shift $\omega$. In prinsiple, we can estimate the source mean size and mean angular position, by computing the normalized cross spectrum, known as coherency, and the phase cross spectrum. Also, we can measure source movements transverse to the viewing direction. For details see Farley et al. (1981) and Providakes (1985).

The Cornell interferometer results for radar aurora revealed a number of interesting features. Providakes et al, (1983) reported measurements showing that type 1 and type 2 echoes may originate from different regions in the scattering volume. Other observations by Providakes et al. (1985) showed that type 3 echoes are associated with horizontal shears in the cross field plasma flow, and with highly localized and irregular structures within the observing volume. Also it was found that type 4 echoes are coming from limited regions in space which led Fejer et al. (1986) to propose the existence of spatially confined regions of elevated

electron temperatures. Haldoupis et al. (1988) reported, with the University of Saskatchewan 50 MHz CW interferometer, that at times high coherence echoes of type 1 or type 3 originate from strongly unstable regions which are related to highly dynamic source regions of transient nature with azimuthal sizes possibly less than 1 km. An example of CW interferometry results is illustrated in Figure 21. This is from an impulsive event whose growth took place during a fraction of a second and whose duration was about 40 s. Figure 21 shows averaged normalized Doppler spectra from both antennas and the corresponding cross spectral estimates. Notice that high coherencies and relatively steady cross phases appear only for the narrow dominant peak of type 3 near -70 Hz. The coherence values in Figure 21 suggest a scattering region with an rms cross sectional area of 2 km or less transverse to the observing direction.



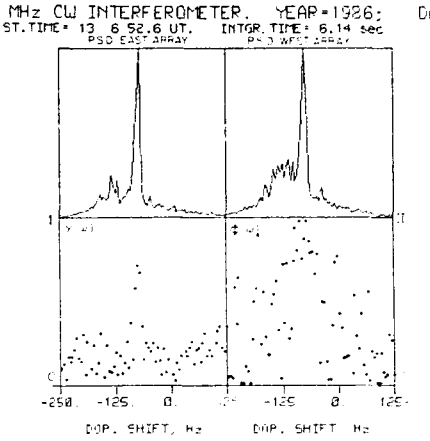




Figure 20. Basic geometry in azimuthal (horizontal) interferometry.

Figure 21. Normalized Doppler spectra (upper two panels) from east and west antennas, and corresponding cross spectral estimates, i.e. the coherency (low left panel) and cross phase spectrum, for a type 3 auroral echo (from the 50 MHz CW interferometer of the radio auroral group at the University of Saskatchewan, Canada).

The existing observations proved that radio auroral interferometry is a promising new technique for investigating dynamics of localized regions of backscatter, and that carefully designed experiments can provide unique altitude and azimuthal information and measure the movements of spatially coherent structures in the auroral plasma. Certainly, we expect to see more of radar auroral interferometer studies in the near future.

## 8. CONCLUDING COMMENTS

In this report, we have reviewed several observations of high latitude auroral E-region irregularities and considered their interpretation in the framework of the existing linear and nonlinear theories. We conclude that, at present, there is a substantial amount of knowledge accumulated through intensive research efforts over the last few years, which has improved the physical picture considerably. On the other hand, our understanding is not yet complete because there are important points that need further study, both theoretical and experimental. Most of the unresolved problems seem to have only auroral identity, i.e. they have no equatorial counterpart. For example, there is now enough evidence to substantiate the claim that there are important differences between 50 MHz equatorial and auroral backscatter processes. These differences focus mainly on the magnetic aspect angle properties of type 1 and 2 echoes and on new echo types observed only in the auroral plasma. Also, the role of altitudes and of electron density in the generation and saturation of instabilities, seem to be much more important in auroral than the equatorial region. Furthermore, it is possible to have in the auroral plasma large amplitude, nearly monochromatic, electron density waves originating from very localized, point-like, regions associated with dynamic sources of transient nature.

The evidence discussed in this paper, emphasizes the need for a general unified theory of radio aurora which would incorporate additional energy sources, beyond those applied in the equator, new (or modified) nonlinear saturation mechanisms and wave energy propagation characteristics. There is now more attention directed to Field Aligned Currents as an important source of energy which can destabilize the plasma. It is not unreasonable to assume that FACs may be the missing element to be included in the equatorial models when applied to aurora However, Chaturvedi et al. (1987) who considered parallel current effects on the E-region instabilities, found that FACs are not enough, when applied in a linearized theory, and some

nonlinearity is necessary to account for results such as backscatter at large magnetic aspect angles.

Besides the necessity for auroral-oriented theoretical models, we also need more carefully designed experim ts to measure with good azimuthal and vertical resolution sheared flows and altitude effects. In particular, the need for precise altitude and callibrated scattering cross section measurements is imminent because there is very few information available on these two important parameters. Furthermore, for more conclusive results it is important to have conducted joint experiments (e.g. using concurrently with auroral radars, rocket and/or incoherent scatter measurements) to carry out correlative studies. These studies would possibly resolve questions concerning the roles of important parameters in the instability processes, such as electron density, elctron density gradients and electron temperatures, and provide useful hints to theorists.

AKNOWLEDGEMENTS. The last few years, there are several people who helped me, in one way or another, to remain active in this high latitude field of research while operating at a mid latitude location. These include Profs G. J. Sofko and J. A. Koehler of the University of Saskatchewan, Canada, Drs E. Nielsen and K. Schlegel of Max-Planck Institute for Aeronomy W. Germany, Prof. A. Egeland of the University of Oslo, Norway, Prof. B. Fejer now at the University of Utah, US, and Prof. J. P. St-Maurice now at the University of Western Ontario, Canada. To all of them I wish to express my gratitude.

REFERENCES

Abel, W. G., and R. E. Newell, Measurements of the afternoon radio aurora at 1295 MHz, J. Geophys. Res., 74, 231, 1969.
André D., The dependence of the relative backscatter cross section of 1-m density fluctuations in the auroral electrojet on the angle between electron drift and radar wave vector, J. Geophys. Res., 88, 8043, 1983.
Bahnsen, A., Ungstrup, C.-G. Falthammar, U. Fahleson, J. K. Olesen, F. Primdhal, F. Spangslev, and A. Pedersen, Electrostatic waves observed in an unstable polar cap ionosphere, J. Geophys. Res., 83, 5191, 1978.
Balsley, B. B., Some characteristics of non-two stream irregularities in the equatorial electrojet, J. Geophys. Res., 74, 2333, 1969.
Balsley, B. B., and D. T. Farley, Radar observations of two dimensional turbulence in the equatorial electrojet J. Geophys. Res., 78, 7174, 1973.
Balsley, B. B., and W. L. Ecklund, VHF power spectra of the radar aurora, J Geophys. Res , 77, 4746, 1972
Bering, E. A., Apparent electrostatic ion cyclotron waves in the diffuse aurora, Geophys. Res. Lett., 80, 4612, 1983.
Bostrom, R., Electrodynamics of the ionosphere, in Cosmical Geophysics, Universitetsforlaget, Oslo, 1973.
Bowles, K. L., R. Cohen, G. R. Ochs, and B. B. Balsley, Radio echoes from field aligned ionization above the magnetic equator and their resemblance to auroral echoes, J. Geophys. Res., 65, 1853, 1960.
Bowles, K. L., B. B. Balsley, and R. Cohen, Field-aligned E region irregularities identified with ion acoustic waves, J. Geophys. Res., 68, 2485, 1963.
Bunemann, O., Excitation of field aligned sound waves by electron streams, Phys. Rev Lett., 10(7), 285, 1963
Cahill, L. J., Jr., R. A. Greenwald, and E. Nielsen, Auroral radar and rocket double-probe observations of electric field across the Harang discontinuity, Geophys. Res. Lett., 5, 687, 1978.
Chaturvedi, P. K., Collisional ion cyclotron waves in the auroral ionosphere, J Geophys. Res., 81,6169, 1976.
Chaturvedi, P. K., J. D Huba, S. L. Ossakow, P. Satyanaaryana, and J. A. Fedder, Parallel current effects on two stream electrojet plasma instabilities, 92, 8700, 1987.
Chesnut, W. G., J. C. Hodges, and R. L. Leadabrand, Auroral backscatter wavelength dependence studies, SRI proj. 5535 Final Rep., Stanford Res. Inst., Menlo Park, Calif., 1968.
Cohen, R , and K L. Bowles, The association of plane-wave electron density irregularities with the equatorial electrojet, J. Geophys. Res., 68, 2503, 1967.
D'Angelo, N., Type III spectra of radio aurora J. Geophys. Res., 78, 3987, 1973.
Dupree, T H., Nonlinear theory of low-frequency instabilities, Phys. Fluids, 11, 2680, 1968
Ecklund, W. L., B. B. Balsley, and R. A. Greenwald, Crossed beam measurements of the diffuse radar aurora, J. Geophys. Res., 80, 1805, 1975.
Farley, D. T , A plasma instability resulting in field-aligned irregularities in the ionosphere, J Geophys. Res., 68, 6083, 1963.
Farley, D. T., Radio wave scattering from the ionosphere, in Methods of Experimental Physics, Vol. 9B, Academic Press, New york, 1971.
Farley, D. T., and B. B. Balsley, Instabilities in the equatorial elctrojet, J. Geophys. Res., 78, 227, 1973
Farley, D. T., and B. G. Fejer, The effect of the gradient drift term on type 1 electrojet irregularities, J. Geophys. Res., 80, 3087, 1975.
Farley, D T., H. M. Ierkic, and B. G. Fejer,, Radar interferometry, Anew technique for studying plasma turbulence in the ionosphere, J. Geophys. Res., 89, 1467, 1981.
Farley, D T., Theory of equatorial electrojet plasma waves: new developments and current s tmos. Terr Phys., 47, 729, 1985.
Fejer, B. G., D. T. Farley, B. B. Balsley, and R. F. Woodman, Vertical structure of the VHF backscattering region in the equatorial electrojet and the gradient drift instability, J. Geophys. Res., 80, 1313, 1975.
Fejer, B. G., and M. G. Kelley, Ionospheric irregularities, Rev. Geophys., 18, 401, 1980.
Fejer, B. G., J. Providakes, and D. T. Farley, Theory of plasma waves in the auroral E region, J. Geophys. Res. Res., 89, 7487, 1984a.
Fejer, B. G., R. W. Reid, D. T. Farley, W. E. Swartz, and M C Kelley, Ion cyclotron waves as a possible source source of resonant auroral E region , J. Geophys. Res., 89, 187, 1984b.
Fejer, B. G., Small scale plasma irregularities in the auroral lower ionosphere, in Proceedings of the 1982-1984 MIT Symposia, Vol. 5, p 73, Scientific Publishers, Cambridge, MA 02139, 1985.
Fejer, B. G., J. Providakes, D. T. Farley, and W. E. Swartz, Auroral Eregion plasma waves and elevated electron temperatures, J. Geophys. Res., 91, 13589, 1986.
Fejer, B. G., and J. Providakes, High latitude E region irregularities: new results, Phys. Scripta, Vol T 18, 167, 1987.
Fejer, J. A. Ionospheric instabilities and fine structure, J. Atmos. Terr. Phys., 41, 895, 1979.

Ferch, R. L. , and R. N. Sudan, Numerical simulations of type II gradient drift irregularities in the equatorial electrojet, J. Geophys. Res., 82, 2283, 1977.
Greenwald, R. A., Diffuse radar aurora and the gradient drift instability, J. Geophys. Res., 79, 4807, 1974.
Greenwald, R. A., and W. L. Ecklund, A new look at radar auroral motions, J. Geophys. Res., 80, 3642, 1975.
Greenwald, R. A., W. L. Ecklund, and B. B. Balsley, Auroral currents, irregularities and luminocity, J. Geophys. Res., 78, 8193, 1973.
Greenwald, R. A., W. L. Ecklund, and B. B. Balsley, Radar observations of auroral electrojet currents, J. Geophys. Res., 80, 3635, 1975.
Greenwald, R. A., W. Weiss, E. Nielsen, and N. R. Thomson, STARE. A new radar auroral backscatter experiment in Northern Scandinavia, Radio Sci., 13, 1021, 1978.
Haldoupis, C. and G. J. Sofko, Doppler spectrum of 42 MHz CW auroral backscatter, Can. J. Phys., 54, 1571, 1976.
Haldoupis, C., and G. J. Sofko, On ion acoustic plasma waves in the radio aurora and the two stream instability, Can. J. Phys., 100, 1978a.
Haldoupis, C. and G. J. Sofko, Short term characteristics of ion acoustic type radio auroral echoes, Can. J. Phys., 56, 292, 1978b.
Haldoupis, C., and G. J. Sofko, VHF double-peaked spectra at negative Doppler shifts in the morning sector of radio aurora, Planet. Space Sci., 27, 233, 1979.
Haldoupis, C., E. Nielsen, and C. K. Goertz, Experimental evidence on the dependence of 140 MHz radar auroral backscatter characteristics on ionospheric conductivity, J. Geophys. Res., 87, 7666, 1982.
Haldoupis, C., and E. Nielsen, Results on relative scattering cross section of 140 MHz auroral backscatter, J. Geophys. Res., 89, 2305, 1984.
Haldoupis, C., E. Nielsen, and H. M. Ierkic, STARE Doppler spectral studies of westward electrojet radar aurora, Planet. Space Sci., 32, 1291, 1984.
Haldoupis, C., P. Hoeq, and E. Nielsen, Doppler spectrum anisotropy observed in 1-m irregularity auroral backscatter, Radio Sci., 20, 676, 1985a.
Haldoupis, C., P. Prikryl, G. J. Sofko, and J. A. Koehler, Evidence for 50 MHz bistatic radio observations of electrostatic ion cyclotron waves in the auroral plasma, J. Geophys. Res., 90, 983, 1985.
Haldoupis, C., G. J. Sofko, and J. A. Koehler, On ion acoustic plasma waves at large magnetic aspect angles in the high latitude E region of the ionosphere, J. Geophys. Res., 91, 5755, 1986.
Haldoupis, C., G. J. Sofko, J. A. Koehler, P. Prikryl, and A. G. McNamara, High resolution Doppler spectrum measurements of 3-m auroral irregularities at large magnetic aspect angles, Phys. Scripta, 35, 910,1987.
Haldoupis, C., G. J. Sofko, J. A. Koehler, P. Prikryl, and M. J. Mckibben, High resolution CW interferometry for auroral ionospheric studies. Preliminary results, J. Geophys. Res., In press, 1988.
Hanuise, C., High - latitude ionospheric irregularities . A review of recent results, Radio Sci., 18, 1093, 1983.
Hofstee, J., and P. A. Forsyth, Ion acoustic waves and aspect sensitivity in radio aurora, J. Atmos. Terr. Phys., 34, 893, 1972.
Ierkic, H. M., Radar observations of the equatorial electrojet irregularities and theory of type 1 turbulence, Ph.D. thesis, Cornell Univ., Ithaca, N. Y., 1980.
Ierkic, H. M., B. G. Fejer, and D. T. Farley, The dependence on zenith angle of the strength of 3-m equatorial electrojet irregularities, Geophys. Res. Lett., 7, 497, 1980.
Igarashi, K., and K. Schlegel, Electron temperature enhancements in the polar E region measured with EISCAT, J. Atmos. Terr. Phys., 49, 273, 1987.
Kamide, Y. R., and A. Brekke, Altitude of the eastward and westward auroral electrojets, J. Geophys. Res., 82, 2851, 1977.
Kelley, M. C., and F. S. Mozer, Electric field and plasma density oscillations due to the high frequency Hall current two stream instability inthe auroral E region, J. Geophys. Res., 78, 2214, 1973.
Keskinen, M. J., R. N. Sudan, and R. L. Ferch, Temporal and spatial power spectrum studies of numerical simulations of type II gradient drift irregularities in the equatorial electrojet, J. Geophys. Res., 84, 1419, 1979.
Kindel, J. M., and C. F Kennel, Topside current instabilities, J. Geophys., Res., 76, 3055, 1971.
Koehler, J. A., G. J. Sofko, and V. Mehta, A statistical study of magnetic aspect effects associated with VHF auroral backscatter, Radio Sci., 20, 689, 1985.
Kudeki E., B. G. Fejer, and D. T. Farley, Interferometer studies of equatorial F region irregularities and drifts, Geophys. Res. Lett., 8, 377, 1981.
Leadabrand, R. L., J. C. Schlobohm, and M. J. Baron, Simultaneous very high frequency and ultrahigh frequency observations of the aurora at Fraserburgh, Scotland, J. Geophys. Res., 70, 4235, 1965.
Lee, K., and C. F. Kennel, Effects of propagation parallel to the magnetic field on the type I electrojet irregularity instability, Planet. Space Sci., 21, 1339, 1973.
Matin, N., and T. B. Jones, Propagation angle dependence of radar auroral E region irregularities, J. Atmos. Terr. Phys., 49, 115, 1987.
McDiarmid, D. R., and McNamara, A. G., VHF radio aurora. simultaneous observation of auroral ionization by two separate radars, Can. J. Phys., 45, 3009, 1967.
McDiarmid., D. R., On the aspect sensitivity of radio aurora, Can. J. Phys., 50, 2558, 1972.
McDonald, B. E., T. P. Coffey, S. L. Ossakow, and R. N. Sudan, Numerical studies of type 2 equatorial electrojet irregularity development, Radio Sci., 10, 247, 1975.
McNamara, A. G., Survey of radio reflections from aurora, in The Radiating Atmosphere , p. 301, Edited by B. M. McCormac, D. Reidel Publishers, Dordrecht, Holland, 1971.
Mitchell, M. J., and J. L. Brown, PAR auroral radar study, vol. V, M and S Computing, Huntsville, Ala., 1976.
Moorcroft, D. R., and R. T. Tsunoda, Rapid scan Doppler velocity maps of the UHF diffuse radar aurora, J. Geophys. Res., 83, 1482, 1978.
Moorcroft, D. R., Dependence of radio aurora at 398 MHz on electron density and electric field, Can. J. Phys., 57, 687, 1979.
Moorcroft, D. R., Propagation of plasma wave energy in the auroral E region, J. Geophys. Res., 89, 2963, 1984.
Moorcroft, D. R., An examination of radio auroral aspect sensitivity, Can. J. Phys., 63, 1005, 1985.
Moorcroft, D. R., and J. M. Ruohoniemi, Nearly simultaneous measurements of radar auroral heights and Doppler velocities at 398 MHz, J. Geophys. Res., 92, 3333, 1987.
Moorcroft, D. R., Estimates of absolute scattering coefficients of radar aurora, J. Geophys. Res., 92, 8723, 1987a.
Moorcroft, D. R., Wavelength dependence of electrostatic plasma wave propagation in the auroral E region, J. Geophys. Res., 92, 2963, 1987b.
Nielsen, E., and K. Schlegel, A first comparison of STARE and EISCAT electron drift velocity measurements, J. Geophys. Res., 88, 5745, 1983.

Nielsen, E., C. Haldoupis, B. G. Fejer, and H. M. Ierkic, Dependence of auroral power spectra variations upon electron drift velocity inthe eastward electrojet, 89, 253, 1985.
Nielsen, E., and K. Schlegel, Coherent radar Doppler measurements and their relationship to the ionospheric electron drift velocity, J. Geophys. Res., 90, 3498, 1985.
Nielsen, E., Aspect angle dependence of mean Doppler velocities of 1-m auroral plasma waves, J. Geophys. Res., 91, 10173, 1987.
Ogawa, T., H. Mori, and S. Miyazaki, Rocket observations of electron density irregularities in the Antarctic E region, J. Geophys. Res., 81, 4013, 1976.
Ogawa, T., B. B. Balsley, W. L. Ecklund, D. A. Carter, and P. E. Johnston, Aspect angle dependence of irregularity phase velocities in the auroral elctrojet, Geophys. Res. Lett., 7, 1081, 1980.
Ogawa, T., H. Mori, s. Miyazaki, and H. Yamagishi, Electrostatic plasma instabilities in highly active aurora observed by sounding rocket S-310JA-7, Mem. Natl. Inst. Polar Res. Spec. Issue Jpn., 18, 312, 1981.
Oksman, J., M. V. Uspensky, G. V. Starkov, G. S. Stepanov, and M. Vallinkoski, The mean fractional electron density fluctuation amplitude derived from auroral backscatter data. J. Atmos. Terr. Phys.,48, 107, 1986.
Ossakow, S. L., k. Papadopoulos, J. Owens, and T. Coffey, Parallel propagation effects onthe type I electrojet instability, J. Geophys. Res., 80, 141, 1975.
Pfaff, R. F., M. C. Kelley, B. G. Fejer, E. Kudeki, C. W. Carlson, A. Pedersen, and B. Hausler, Electric field and plasma density measurements in the auroral electrojet, J. Geophys. Res., 89, 236, 1984.
Prikryl, P., J. A. Koehler, G. J. Sofko, D. J. McEwen, and D. Steele, ionospheric ion cyclotron wave generation inferred from coordinated Doppler radar, optical and magnetic observations, J. Geophys. Res., 92, 3315, 1987.
Primdahl, F., and A. Bahnsen, Auroral E-region diagnosis of non-linearly stabilized plasma waves, *Ann. Geophys.*, Geophys., 3, 57, 1985..
Primdahl, F., Polar ionospheric E region plasma wave stabilization and electron heating by wave-induced enhancement of the electron collision frequency, Physica Scripta, 33, 187, 1986.
Providakes, J. F., W. E. Swartz, D. T. Farley, B. G. Fejer, First VHF auroral radar interferometer observations, Geophys. Res. Lett., 10, 401, 1983.
Providakes, J. F.,Radar interferometer observations and theory of plasma irregularities in the auroral ionosphere, Ph.D. thesis, Cornell Univ., Ithaca, N. Y., *1985.*
Providakes, J. F., D. T. Farley, W. E. Swartz, and D. Riggin, Plasma irregularities associated with morning discrete auroral arcs . Radar interferometer observations and theory, J. Geophys. Res., 90, 7513, 1985.
Riggin, D., W. E. Swartz, J. Providakes, and D. T. Farley, Radar studies of long wavelength waves associated with mid-latitude sporadic E layers, J. Geophys. Res., 91, 8011, 1986.
Reed, R. W., Spectra from 50 MHz auroral radar echoes, Ph.D. thesis, Cornell Univ., Ithaca, N.Y., 1980.
Robinson, T. R., Towards a self-consistent non linear theory of radar auroral backscatter, J. Atmos. Terr. Phys., 48, 417, 1986.
Rogister, A., and N. D Angelo, Type 2 irregularities in the equatorial electrojet, J. Geophys. Res., 75, 3879, 1970.
*Ruohoniemi*, J. M., and D. R. Moorcroft, Radar auroral echo heights as seen by a 398 MHz phased array radar, Radio Sci., 20, 719, 1985.
Sato, T., Unified theory of type I and II irregularities in the equatorial electrojet, J. Geophys, Res., 78, 2232, 1973.
Schlegel, K., and *J. P. St-Maurice, Anomalous heating of the polar E region by unstable plasma waves.* 1. Observations, J. Geophys. Res., 86, 1447, 1981.
Schlegel, K., and J. P. St-Maurice, Note on the parallel propagation effects of unstable Farley-Bunemann waves at high latitudes, Planet. Space Sci., 30, 315, 1982.
Schlegel, k., Interpretation of auroral radar experiments using a kinetic theory of the two stream instability, Radio Sci., 18, 108, 1983.
Schlegel, K., E. C. Thomas, and D. Ridge, A statistical study of auroral radar spectra obtained with SABRE, *J. Geophys. Res., 91, 13483, 1986.*
Sinclair, I. P. W., and P. A. Forsyth, Radio auroral scattering at two closely spaced frequencies, Can. J. Phys., 58, 1485, 1980.
Siren, J. C., J. R. Doupnik, and W. L. Ecklund, A comparison of auroral currents measured by the Chatanika radar with 50 MHz backscatter observed from Anchorage, J. Geophys. Res., 82, 3577, 1977.
Schunk, R. W., and A. F. Nagy, Electron temperatures in the F region of the ionosphere. Theory and observations, Planet. Space Sci., 21, 1875, 1978.
Sofko, G. J., J. A. Koehler, J. Gimler, A. G. McNamara, and D. R. McDiarmid, Radio aurora magnetic and streaming aspect sensitivities on 6 simultaneous links at 50 MHz, Adv. Space Sci., 2(7), 149, 1983.
Sofko, G. J., and A. Kavadas, Periodic fading in 42 MHz auroral backscatter, J. Geophys. Res. 74, 3651, 1969.
Sofko, G. J., and A. Kavadas, Polarization characteristics of 42 MHz auroral backscatter, J. Geophys. Res., 76, 1778, 1971.
Starkov, G. V., J. Oksman, M. V. Uspensky, and A. V. Kustov, On the dependence of radar aurora amplitude on ionospheric electron density, J. Geophys., 52, 49, 1983.
St-Maurice, J. P., K. Schlegel, and P. M. Banks, Anomalous heating of the polar E region by unstable plasma waves , 2. Theory. J. Geophys. Res., 86, 1453, 1981.
St-Maurice, J. P., *A nonlocal theory of high latitude Farley-Bunemann instability*, J. Geophys. Res., 90, 5211, 1985.
St-Maurice, J. P., A unified theory of anomalous resistivity and Joule heating effects in the presence of ionospheric E region irregularities, J. Geophys. Res., submitted, 1987.
Sudan, R. N., J. Akinrimisi, and D. T. Farley, Generation of small scale irregularities in the equatorial electrojet, J. Geophys. Res., 78, 240, 1973.
Sudan, R. N., Unified theory of type 1 and type 2 irregularities in the equatorial electrojet, J. Geophys. Res., 1983.
Tsunoda, R. T., Electric field measurements above a radar scattering volume producing diffuse auroral echoes, J. Geophys. Res., 80, 4297, 1975.
Tsunoda, R. T. Doppler velocity maps of the diffuse radar aurora, J. Geophys. Res., 81, 425, 1976.
Tsunoda, R. T., and Presnell, On the threshold electric field associated with the 398 MHz diffuse radar aurora, J. Geophys. Res., 81, 88, 1976.
Timofeev, E., and Y. Miroshnikov, Altitude characteristics of radar aurora as seen by a 90 MHz double-altitude radar system operated at Karmaselga, Karelia, J. Geophys., 51, 1982.
Unwin, P. S., The radio aurora, Ann. Geophys., t 28, fasc. 1, p 111, 1972.

Unwin, R. S., and P. V. Johnston, Height dependence in the power spectrum of diffuse radar aurora, J. Geophys. Res., 86, 5733, 1981.
Uspensky, M. V., R. J. Pellinen, W. Baumjohann, G. V. Starkov, E. Nielsen, G. Sofko, and K. U. Kaila, Spatial variations of ionospheric conductivity and radar auroral amplitude in the eastward electrojet region during pre-substorm conditions, J. Geophys., 52, 40, 1983.
Volocevich, A. V., and V. A. Liperovskiy, Generation of small scale inhomogeneities in a turbulent plasma and radio auroras, Geomagn. Aeron., 14, 58, 1975.
Villain, J. P., R. A. Greenwald, K. B. Baker, and J. M. Ruohoniemi, HF radar observations of E region plasma irregularities produced by oblique electron streaming, J. Geophys. Res., 92, 12327, 1987.
Wang, T. N. C., and R. T. Tsunoda, On a crossed field two stream plasma instability in the auroral plasma, J. Geophys. Res., 80, 2172, 1975.

## DISCUSSION

**J.Fejer**

There have been suggestions on theoretical grounds that powerful high frequency transmissions, transmitted from ionospheric modification facilities, could modify the natural instability mechanisms. Is there any experimental evidence for this in the $E$ region?

**Author's Reply**

One may not expect a detectable effect during disturbed periods. On the other hand, this would be possible during quiet periods at low $Kp$ activity. It has been observed, during simultaneous operation of the HF Heater Facility in Trömso and the STARE radar, that weak echoes were received during times when the heater was on. This, however, if I am not mistaken, was observed during periods when naturally produced echoes were absent.

# IONOSPHERIC IRREGULARITIES AND THEIR EFFECTS ON ELECTROMAGNETIC WAVES PROPAGATING THROUGH THEM

K. C. Yeh

Department of Electrical and Computer Engineering, University of Illinois
1406 W. Green Street, Urbana, IL 61801 USA

## SUMMARY

Global sounding and probing of the ionosphere over the last several decades have revealed the frequent presence of irregularities throughout various ionospheric regions. Using the collected data, many comprehensive studies have been made and reviewed by various authors. Models have been constructed and continuously improved with additional data. Extensive morphological studies clearly indicate a strong dependence on the phase of solar cycle, the geomagnetic latitude, and the local time. Data have also been collected that enable us to characterize the electron density fluctuations statistically in their space-time behavior. Such characterizations are necessary in connecting medium properties to wave properties.

The ionosphere is known to be a dispersive medium. In the presence of irregularities, severe scattering may additionally take place. A radio wave propagating through such an environment may therefore experience fading and scintillation. The strength of a scintillating signal is usually measured by its scintillation index $S_4$, defined as the normalized root-mean-square intensity fluctuations above the mean. The dependence of $S_4$ on radio frequency has been investigated by using multifrequency data. The results are shown to agree with theory. The higher-order statistics that are used to describe the space-time signal have also been measured. Using spaced antennas, a cross correlation function can be calculated in addition to $S_4$ and self spectrum. Several methods exist that connect the cross correlation function to the ionospheric drift velocity and the velocity fluctuations. Several days of data obtained at an equatorial station will be shown as examples. They all show a continuous decay in velocity fluctuations throughout the evening with a tendency to rise near the boundary of an irregularity patch.

## 1. INTRODUCTION

The Earth's ionosphere is under continuous agitation. Even in the absence of external forces, the thermal agitation alone can produce fluctuations in electron density. These thermally excited density fluctuations are very weak and are detectable only with very powerful radars, commonly called incoherent-scatter radars [Gordon, 1958; Farley, 1971; Evans, 1975]. Frequently, the ionosphere is under the influence of external forces in addition to the perpetual gravity. Among them are the electric field, Lorentz force, neutral winds, strong gradients, gravity waves and tides. Some of these forces can make the ionospheric plasma unstable. Under such conditions an initial perturbation can grow rapidly to very large values until limited by nonlinear processes. As a result, the plasma thermal fluctuations are enormously amplified. For example, it has been reported that the radar scatter from thermally excited fluctuations at the magnetic equator can surge by 80 dB in a matter of minutes [McClure and Woodman, 1972]. This corresponds to an increase of eight decades in spectral intensity at one-half the radar wavelength, a tremendous increase indeed. No wonder, historically the detection of ionospheric density fluctuations, also known as ionospheric irregularities, was made under such strong perturbations. This happened as early as 1938 [Booker and Wells, 1938] soon after the introduction of frequency-sweeping ionosondes. The otherwise clean echo traces from the F region were found to spread and become blurred. These blurred echoes have been commonly called the spread F echoes [Herman, 1966]. A historical account of these events can be found elsewhere [Special Issue, 1974].

One experimental method in the investigation of ionospheric irregularities is the use of a transionospheric radio beacon. A radio wave passing through a perturbed ionosphere will sense the perturbation of the dielectric permittivity $\varepsilon_1$ given by

$$\varepsilon_1(\vec{r},t) = -\frac{(f_{po}/f)^2}{1-(f_{po}/f)^2}\,\frac{\Delta N(\vec{r},t)}{N_o} \tag{1}$$

where $\Delta N$ is the space-time perturbation in electron density $N_o$ is the background electron density with a corresponding plasma frequency $f_{po}$ and f is the radio frequency. Because of the dispersive nature of the ionosphere the fluctuating part of the relative dielectric permittivity is frequency dependent. For most radio beacon experiments the radio frequency is much larger than the background plasma frequency. Even under this condition, $\varepsilon_1$ is inversely proportional to the square of f. This immediately suggests investigations by a multifrequency beacon over a band within which the lowest frequency will be most sensitive to density fluctuations while the highest frequency will be least sensitive.

Radio beacon investigations of ionospheric irregularities rely on perturbations imparted on wave parameters, such as amplitude or phase, by scattering from irregularities. The severity of fluctuations in wave parameters, alternately known as scintillations, is an indication of irregularity intensities plus a number of other parameters. The link is provided by the scintillation theory, which shows that, in addition to irregularity strength, the severity of scintillations also depends on the radio frequency, the irregularity shape and propagation geometry. Some of these results are reviewed in section 2. The scintillating signals can be spectrally analyzed. For weak scintillations, the phase spectrum follows a power law while the amplitude spectrum is nearly flat to a Fresnel frequency above which a roll-off to a power law behavior can be expected. The use of such information to study the ionospheric irregularity is reviewed in section 3. The spatial structure of the received wave can be measured by spaced antennas. Their use can determine heights, drifts and random velocities as discussed in section 4. Finally, this paper is concluded in section 5.

## 2. SCINTILLATION INDICES

The easiest parameter to measure experimentally is the amplitude. In the presence of strong ionospheric irregularities, the amplitude is found to fluctuate or scintillate. Several quantities [Briggs and Parkin, 1963] can be defined to measure the severity of fluctuations. These quantities are called the scintillation indices. The second moment of the complex amplitude u is the ensemble averaged intensity, $\langle I\rangle = \langle uu^*\rangle$. Under the forward scatter approximation, $\langle I\rangle$ is found to be a constant, i.e. independent of $\Delta N$ and propagation geometry. One is thus forced to go to the fourth moment. The corresponding scintillation is popularly denoted by $S_4$ which is defined by

$$S_4^2 = (\langle I^2\rangle - \langle I\rangle^2)/\langle I\rangle^2 \tag{2}$$

The phase scintillation is more difficult to measure. What is done on radio beacon experiments is to transmit several harmonically related frequencies whose dispersive phases are processed using the highest frequency as a reference. The phase fluctuations $\phi$ at a frequency lower than the reference frequency is used to compute the root-mean-square value $\sigma_\phi$. It has been found that the computation of $\sigma_\phi$ depends on the cutoff time of the detrending filter. To facilitate the intercomparison of data, this cutoff time should also be specified. For orbiting satellites, a 10 s cutoff time is usually adopted.

Theoretically, the dependence of $S_4$ on the standard deviation of fluctuating electron density, $\sigma_{\Delta N}$, can be computed using the multiple scattering scintillation theory [Yeh and Liu, 1982]. It has been found that for a given frequency and a given propagation geometry, $S_4$ increases linearly with $\sigma_{\Delta N}$ as predicted by the Rytov solution when $\sigma_{\Delta N}$ is small (See Fig. 1). The rate of increase in $S_4$ slows with further increase in $\sigma_{\Delta N}$ and eventually saturates at $S_4 = 1$ for very large $\sigma_{\Delta N}$. When $S_4 \approx 1$, the amplitude has a Rayleigh distribution. The asymptotic approach to the saturation value of $S_4 \approx 1$ can come from below or from above, depending on the spectral index and the size of the Fresnel scale relative to the outer scale. The details of this problem have been summarized by Rino [1987].

Under weak scintillation conditions, the scintillation index $S_4$ has a frequency dependence given below [Jokippi and Hollweg, 1970].

$$S_4 \propto f^{-n} \tag{3}$$

When the irregularity spectrum is of the form $\Phi_{\Delta N}(k) \propto \kappa^{-p}$, the frequency exponent n in eq. (3) is linked to the spectral index p by $n = (2+p)/4$. Thus processing of scintillation data for $S_4$ at two or more frequencies can provide the value of p, provided that the values of $S_4$ do not exceed about 0.1 on all frequencies. As $S_4$ values increase, multiple scattering becomes important. Its effect is to weaken the dependence on $f$ predicted by eq. (3). Eventually under saturation conditions, $S_4$ loses all its frequency dependence. The frequency scaling law of eq. (3) under the weak scintillation condition and the loss of frequency dependence under the saturation condition are clearly revealed in the theoretical result of Fig. 1. Experimentally, this result can be very compactly presented when it is realized that $S_4$ in a power-law environment depends only on $S_4^{(o)}$ irregularity slab thickness and the spectral index p [Rino and Liu, 1982]. Here $S_4^{(o)}$ is the scintillation index predicted by the Rytov solution. The dependence of $S_4$ on the propagation geometry and the root-mean-square density fluctuations is completely absorbed by its dependence on $S_4^{(o)}$. In this regard, $S_4^{(o)}$ can be viewed as a universal parameter. In a multifrequency scintillation experiment, the highest frequency invariably satisfies the weak scatter condition. The frequency scaling law of eq. (3) can be used to predict what $S_4^{(o)}$ should be at a lower frequency. At this same lower frequency $S_4$ itself can be determined from experimental data. Thus a plot of $S_4$ vs $S_4^{(o)}$ can be made. This was done by Rino and Liu [1982]. Their results are reproduced in Fig. 2. The points follow a straight line for weak scintillations and eventually saturate to $S_4 = 1$ when scattering is very strong. There is also some scatter of points which may arise because the irregularities are not necessarily isotropic, the outer scale may become comparable to the Fresnel scale, the spectral index may depart from $p = 4$ and the determination of $S_4$ may be subject to some experimental errors. Furthermore, some data points seem to suggest the existence of focussing effects when $S_4$ rises to about unity.

The scintillation also depends on the geometric shape of the irregularities. This problem has been studied thoroughly by Singleton [1970]. Observationally, a slight broadening of irregularities in the magnetic meridian plane has been reported at high latitudes [Singleton, 1973; Martin and Aarons, 1977]. At Fairbanks, Alaska, Fremouw et al. [1977] observed a persistent enhancement in intensity and phase scintillations when the Wideband satellite was in the magnetic meridian plane. This is dramatically represented in Fig. 3. Such aspect-sensitive observations allow Fremouw and Lansinger [1981] to conclude that most irregularities in the auroral region are rod-like while sheet-like irregularities are confined to the nightside of the high-latitude ionosphere.

Global observations of the scintillation data have been used to investigate the scintillation morphology and its relation to other geophysical events, e.g., [Yeh and Swenson, 1964; Aarons, 1982]. There have also been periodic attempts to model this phenomenon empirically using an ever larger data base [Fremouw and Secan, 1984]. Pictorially the global situation can be sketched as shown in Fig. 4 [Basu et al., 1988]. This figure shows only the worst case in the year of a given solar epoch. Not shown are the details such as the interesting seasonal behavior especially at the equator.

## 3. SECOND ORDER STATISTICS

The second order statistics of a scintillating signal observed at a single location is represented by its spectrum or correlation function. For a one-component power-law irregularity spectrum under weak scintillation conditions, the theory predicts a flat amplitude spectrum for low spatial frequencies, $\kappa_x$, up to the Fresnel wave number $\kappa_F$ given by

$$\kappa_F = (\pi k/z)^{1/2} \tag{4}$$

where k is the wave number of the radio wave and z is the slant distance from the receiver to the middle of the irregularity slab. For $\kappa_x$ above $\kappa_F$, the amplitude rolls off in the form $\kappa_x^{1-p}$ where p is the spectral index of irregularities. Additionally, there may exist deep spectral nulls at $\kappa_m$ given by

$$\kappa_m = \sqrt{2m}\ \kappa_F\ , \quad m = 1, 2, 3.... \tag{5}$$

These nulls arise from Fresnel oscillations and are usually smeared out unless the irregularity slab is very thin [Liu and Yeh, 1977].

On the other hand, the low spatial frequency part of the phase spectrum is not affected by Fresnel filtering but by the detrending process in the data analysis and by the finite data length. Thus the phase spectrum will ideally decay like $\kappa_x^{1-p}$. If the irregularity slab is thin, Fresnel nulls on the phase spectrum may appear at

$$\kappa_m = \sqrt{2m-1}\ \kappa_F\ , \quad m = 1, 2, 3,.... \tag{6}$$

Observationally, the scintillation signal is recorded as a function of time. The processing of the time series gives the temporal spectrum as a function of frequency $\nu$. If the temporal fluctuations in the received wave are produced by a frozen flow of velocity $v_x$, the temporal variations and spatial variations can be related. For spectra the proper transformation is $\nu = \kappa_x v_x/2\pi$. Hence for weak amplitude scintillations, the temporal spectrum in a log-log plot is supposed to show a corner frequency at

$$\nu_F = \kappa_F v_x/2\pi \tag{7}$$

If the corner frequency $v_F$ is determined from the amplitude spectrum and the drift velocity $v_x$ is obtained by, for example, spaced receiver experiment, eqs. (4) and (7) can be used to determine the slant distance z to the middle of the irregularity slab as was done [Yeh et al., 1981].

Several studies have been made on the inferred behavior of spectral index p deduced from asymptotic spectral slopes of the scintillation data. For example, Basu et al. [1980] have found that at the equator the spectra are fairly shallow in the first hour after the onset of scintillation. Thereafter, the spectra tend to steepen with time, suggesting progressive decay of small irregularities. This kind of behavior has been confirmed by observations of Spatz et al. [1988] who also obtain a curious notch centered at about 01:00 LT. This is shown in Fig. 5. At equatorial latitudes, a two-component power-law irregularity spectral behavior has been frequently observed [Rino et al., 1981; Kelley and McClure, 1981]. Their existence makes the scintillation data interpretation more difficult, especially when scintillation is strong. In this case, considerable broadening beyond Fresnel frequency $v_F$ can be expected extending to a new corner frequency, above which an initial steep Gaussian-like roll-off is followed by an eventual power-law decay. Such a behavior is supported by computer simulations [Franke and Liu, 1983; 1985].

As mentioned earlier, the Fresnel nulls are usually height smeared throughout the irregularity slab [Liu and Yeh, 1977]. However, when the slab is thin, these nulls can be very distinct as occasionally being observed. For the bottomside sinusoidal irregularities in the equatorial F region, these Fresnel nulls up to the seven orders have been observed [Basu et al., 1986a]. One proposed explanation is that such irregularities are concentrated in a layer not exceeding 50 km.

Alternate to spectral analysis, the scintillating amplitude can be processed for autocorrelation functions. The time lag for the normalized autocorrelation function to decrease to 0.5 can be defined as the correlation interval. It has been found that, in the multiple scattering regime, the correlation interval increases with frequency because the decorrelation brought about by multiple scattering decreases [Umeki et al., 1977]. On the other hand, in the weak scatter regime, the correlation interval is controlled by Fresnel filtering indicated by $\kappa_F$ of eq. (4) resulting in a decrease in the correlation interval as a function of frequency. Therefore, as a function of frequency, the correlation interval will increase until the multiple scattering is much weakened. Thereafter it is followed by a decreasing correlation interval when the frequency is high enough that the weak scatter theory is valid. Experimental support to this phenomenon has been found [Umeki et al., 1977; Yeh and Liu, 1982].

## 4. SPATIAL STRUCTURES

The spatial structures of the scintillation pattern on the ground can be investigated by placing more than one receiver there. Experimental results show that the pattern evolves with time as it drifts. Several methods have been developed to measure the variations of the drifting pattern [Briggs et al., 1950]. Central in these methods are two velocities: the drift velocity $V_0$ and the characteristic velocity $V_c$. These methods have been applied to the scintillation data [Wernik et al., 1983] and tested for internal consistency [Vacchione et al., 1987]. It has been found that Briggs' peak-value method yields the most consistent results. Applying this method to the scintillation data recorded on April 23, 1983 at Guam [Spatz et al., 1988], the drift velocity is found to decrease from about 150 m/s to 70 m/s in a matter of several hours as shown in Fig. 6. These values are within the range of values obtained elsewhere [Yeh et al., 1981; Basu et al., 1986b]. It should be cautioned that the apparent drift along the east-west baseline is caused by the projection of the vector velocity along the baseline, which has both the vertical and horizontal components. Thus, the drift velocity can be interpreted as the zonal velocity only when the vertical movements of the ionosphere are negligible.

Another parameter that can be deduced from the spaced receiver data is $V_c$. Originally in Briggs' work, $V_c$ is used to characterize the space-time decorrelation speed as the scintillation pattern evolves with time. The more recent work [Wernik et al., 1983; Vacchione et al., 1987; Spatz et al., 1988] has related $V_c$ to the rms velocity fluctuations $\sigma_v$ in a locally frozen flow. Using the peak-value method $V_c$ for the same data as shown in Fig. 6 has been computed and depicted in Fig. 7. Also shown in Fig. 7 are values of $S_4$ index. This figure demonstrates that for the first hour and a half after the onset of scintillation, $V_c$ can be as large as 50 m/s which is about one-third of $V_0$. Thereafter, $V_c$ decays rapidly to 10 m/s or lower. There is also clear indication that $V_c$ increases to almost twice its value when the western or the eastern wall of the irregularity patch is approached. The decay of $V_c$ with time seems to be the dominant feature at the equator. A mass plot of $V_c$ for eight nights is given in Fig. 8. The decay of $V_c$ with time is obvious even though the rise of $V_c$ near the wall of a scintillation patch is smeared out.

## 5. CONCLUSIONS

A review of the scintillation method applied to the investigation of ionospheric irregularities has been made. Interestingly, the suggestion of ionospheric irregularities to have power-law spectra first came from scintillation measurements [Procello and Hughes, 1969; Elkins and Papagiannis, 1969; Rufenach, 1972]. Since then the scintillation measurements have been expanded and made in conjunction with other experiments to enhance our understanding of geomorphology of ionospheric irregularities. Such an understanding has a direct application to satellite communications.

## 6. ACKNOWLEDGMENT

The work was supported by the National Science Foundation under a grant number ATM 84-14134.

## REFERENCES

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

Basu, S., Su. Basu, C. E. Valladares, A. DasGupta, and H. E. Whitney, Scintillations associated with bottomside sinusoidal irregularities in the equatorial F region, *J. Geophys. Res., 91*, 270-276, 1986a.

Basu, S., Su. Basu, J. LaBelle, E. Kudeki, B. G. Fejer, M. C. Kelley, H. E. Whitney, and A. Bushby, Gigahertz scintillations and spaced receiver drift measurements during project Condor equatorial F region rocket campaign in Peru, *J. Geophys. Res., 91*, 5526-5538, 1986b.

Basu, S., E. MacKenzie, and Su. Basu, Ionospheric constraints on VHF/UHF communications links during solar maximum and minimum periods. To appear in *Radio Sci.*, 1988.

Basu, S., J. P. McClure, S. Basu, W. B. Hanson, and J. Aarons, Coordinated study of equatorial scintillation and in situ and radar observations of nighttime F region irregularities, *J. Geophys. Res., 85*, 5119-5130, 1980.

Booker, H. G., and H. W. Wells, Scattering of radio waves by the F-region of the ionosphere, *J. Geophys. Res., 43*, 249-256, 1938.
Briggs, B. H., and I. A. Parkin, On the variation of radio star and satellite scintillations with zenith angle, *J. Atmos. Terr. Phys., 25*, 339-366, 1963.
Elkins, T. J., and M. D. Papagiannis, Measurement and interpretation of power spectrums of ionospheric scintillation at sub-auroral location, *J. Geophys. Res., 74*, 4105-4115, 1969.
Evans, J. V., High power radar studies of the ionosphere, *Proc. IEEE, 63* , 1636-1650, 1975.
Farley, D. T., Radio wave scattering from the ionosphere, in *Methods of Experimental Physics*, edited by R. H. Lovberg and H. R. Grien, Academic Press, New York, 9B. Ch. 14, 1971.
Franke, S. J., and C. H. Liu, Observations and modeling of multi-frequency VHF and GHz scintillations in the equatorial region, *J. Geophys. Res., 88(A9)*, 7075-7085, 1983.
Franke, S. J., and C. H. Liu, Modeling of equatorial multifrequency scintillation, *Radio Sci., 20*, 403-415, 1985.
Fremouw, E. J., and J. M. Lansinger, Dominant configurations of scintillation-producing irregularities in the auroral zone, *J. Geophys. Res., 86*, 10,087-10,093, 1981.
Fremouw, E. J., C. L. Rino, R. C. Livingston, and M. C. Cousins, A persistent subauroral scintillation enhancement observed in Alaska, *Geophys. Res. Lett., 4*, 539-542, 1977.
Gordon, W. E., Incoherent scattering of radio waves by free electrons with applications to space exploration by radar, *Proc. IRE, 46*, 1824-1829, 1958.
Herman, J. R., Spread F and ionospheric F region irregularities, *Rev. Geophys., 4*, 255-299, 1966.
Jokippi, J. R., and J. V. Hollweg, Interplanetary scintillation and the structure of solar wind fluctuations, *Astrophys. J., 160*, 745-753, 1970.
Kelley, M. C., and J. P. McClure, Equatorial spread-F: A review of recent experimental results, *J. Atmos. Terr. Phys., 43*, 427-436, 1981.
Liu, C. H., and K. C. Yeh, Model computations of power spectra for ionospheric scintillations at GHz frequencies, *J. Atmos. Terr. Phys., 39*, 149-156, 1977.
Martin, E., and J. Aarons, F layer scintillations and the aurora, *J. Geophys. Res., 82*, 2717-2722, 1977.
McClure, J. P., and R. F. Woodman, Radar observations of equatorial spread-F in region of electrostatic turbulence, *J. Geophys. Res., 77*, 5617-5671, 1972.
Procello, L. J., and L. R. Hughes, Observed fine structure of a phase perturbation induced during transauroral propagation, *J. Geophys. Res., 73*, 4105-4115, 1969.
Rino, C. L., The applications of strong-scatter theory to transionospheric radio propagation, in multiple scattering of waves in random media and random rough surfaces, edited by V. V. Varadan and V. K. Varadan, 311-326, The Pennsylvania State University, 1987.
Rino, C. L., and C. H. Liu, Intensity scintillation parameters for characterizing transionospheric radio signals, *Radio Sci, 17*, 279-284, 1982
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*, 2411-2420, 1981.
Rufenach, C. L., Power law wave number spectrum deduced from ionospheric scintillation observations, *J. Geophys. Res., 77*, 4761-4772, 1972.
Singleton, D. G., The effect of irregularity shape on radio star and satellite scintillations, *J. Atmos. Terr. Phys., 32*, 315-343, 1970.
Singleton, D. G., The dependence of high-latitude ionospheric scintillation on zenith angle and azimuth, *J. Atmos. Terr. Phys., 35*, 2253-2265, 1973.
Spatz, D. E., S. J. Franke, and K. C. Yeh, Analysis and interpretation of spaced receiver scintillation data recorded at an equatorial station, To appear in *Radio Sci.*, 1988.
Special Issue: Fifty Years of the Ionosphere, *J. Atmos. Terr. Phys., 36*, No. 12, 1974.
Umeki, R., C. H. Liu, and K. C. Yeh, Multifrequency spectra of ionospheric amplitude scintillations, *J. Geophys. Res., 82*, 2752-2760, 1977.
Vacchione, J. D., S. J. Franke, and K. C. Yeh, A new analysis technique for estimating zonal irregularity drifts and variability in the equatorial F region using spaced receiver scintillation data, *Radio Sci., 22*, 745-756, 1987.
Wernik, A. W., C. H. Liu, and K. C. Yeh, Modeling of spaced-receiver scintillation measurements, *Radio Sci., 18*, 743-764, 1983.
Yeh, K. C., J. P. Mullen, J. R. Medeiros, R. F. daSilva, and R T Medeiros, Ionospheric scintillation observations at Natal, *J. Atmos. Terr. Phys., 86*, 7527-7532, 1981.
Yeh, K. C., and C. H. Liu, Radio wave scintillation in the ionosphere, *Proc. IEEE, 70*, 324-360, 1982.
Yeh, K. C., and G. W. Swenson, Jr., F-region irregularities studied by scintillation of signals from satellites, *J. Res. Nat. Bur. Stand. (Radio Sci.), 68D*, 881-894, 1964.

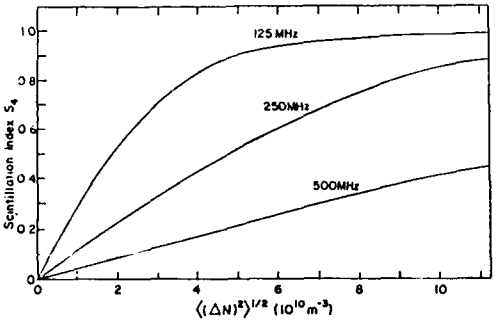



Fig. 1. Scintillation index $S_4$ as a function of $\langle(\Delta N)^2\rangle^{1/2}$ at 125, 250 and 500 MHz. The slab of irregularities has a thickness 50 km and is 237.5 km from the receiver. Computations are made for a power-law irregularity spectrum with an outer scale of 500 m and a spectral index of 4. The background electron density is $10^{12}$ m$^{-3}$. [After Yeh and Liu, 1982].

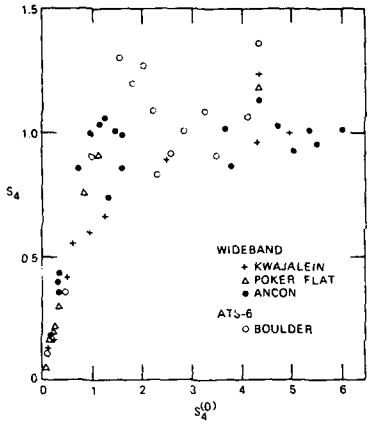



Fig. 2. Multistation, multisatellite and multifrequency scintillation data. The scintillation index $S_4$ is calculated from experimental data according to formula (2). The scintillation index $S_4^{(0)}$ is calculated from the observed data at a highest frequency at which the weak scatter theory is valid and scaled to a lower frequency by using the scaling law (3) with $n$ = 1.5. [After Rino and Liu, 1982].

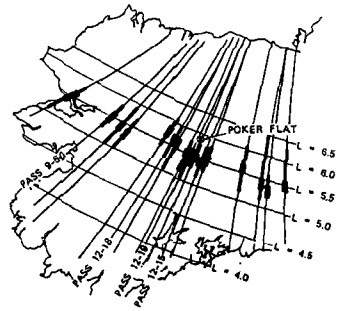



Fig. 3. Map of Alaska showing satellite tracks projected to an ionospheric height of 350 km for different passes. The track is thickened in the region when enhanced scintillations were observed. [After Fremouw et al., 1977].

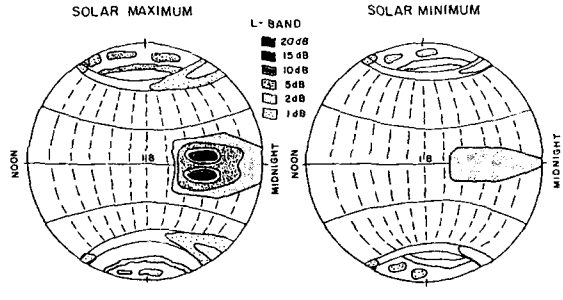



Fig. 4. Global behavior of scintillation fades at L band during the solar maximum and solar minimum conditions. What is shown is the worst case for a given solar epoch. [After Basu et al., 1988].

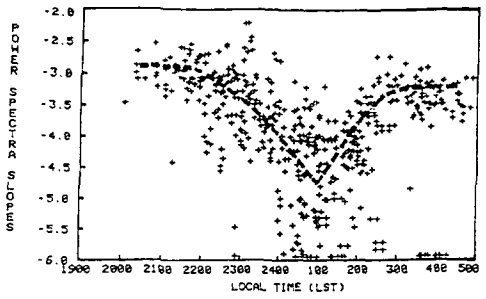



Fig. 5. Measured amplitude power spectra slopes for 20 nights with $S_4 < 0.5$ at a frequency of 250.55 MHz. Data were recorded at Guam (12.4°N, 147.0°E) during 1982 through 1983 of signals transmitted by the Pacific Ocean Fleet Satellite (0.00°N, 176.50°E). [After Spatz et al., 1988].

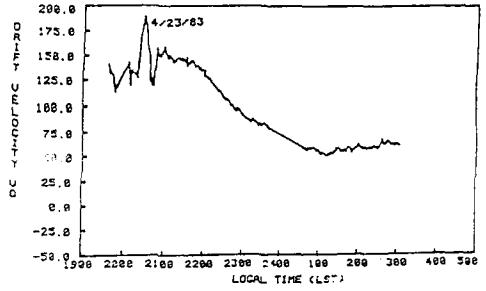



Fig. 6. Drift velocity deduced from the spaced scintillation data using the peak-value method for April 23, 1983 at Guam.

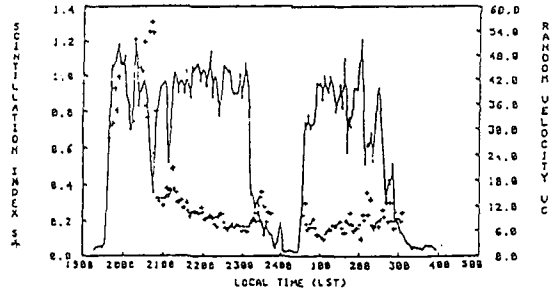



Fig. 7. The scintillation index $S_4$ (solid curve, left scale) and the random velocity $V_c$ (crosses, right scale) deduced from the same data as shown in Fig. 6.

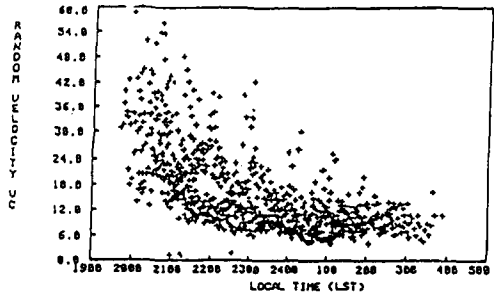



Fig. 8. Mass plot of $V_c$ for eight nights for the data recorded at Guam.

## DISCUSSION

**H.Rishbeth**

Does there exist any general theory of the origin of ionospheric irregularities at mid-latitudes?

**Author's Reply**

As reminded by Dr. J. Fejer, there was a theory proposed by Perkins in a paper published in JGR in the 1960's or early 1970's. However, the field of mid-latitude irregularities has not received much attention in recent years.

**E.Szuszczewicz**

Please give us some guidelines on what you would think are quantitative thresholds for power spectral index and frequency domains most likely to affect different propagation channels going upwards from the HF to the gigahertz domain.

**Author's Reply**

In the review paper by Kelly, he has identified four regimes and he calls them the low, intermediate, transition and high frequency regimes using different wave numbers. As far as the scintillation is concerned, the important regime is near the Fresnel zone. That depends on the frequency we are using, but would be of the order of a kilometer in that region.

# ELECTROSTATIC TURBULENCE IN THE HIGH LATITUDE IONOSPHERE

by

J.-C.Cerisier, H.Mounir and J. J.Berthelier
*Centre de Recherches en Physique de l'Environnement*
4, Avenue de Neptune
94107 Saint Maur Cedex
France

## ABSTRACT

Two components of the Extremely Low Frequency (ELF) electric field are measured on board the low altitude (400—2000 km) polar orbiting satellite Aureol -3. At high latitude, the main wave-field component observed in the frequency range up to a few hundred Hertz is the electrostatic turbulence.

The frequency spectrum measured in the satellite reference frame usually follows a power law ($E^2 \propto E_0 f^{-\alpha}$), thus allowing to describe the turbulence with two parameters (amplitude and spectral index). These characteristic parameters will be discussed in relation with the possible free energy sources: field-aligned current, particle precipitation, electron density gradient.

The relation between the frequency spectrum and the wave number spectrum is not unambiguous. However, under the hypothesis of time-independent turbulence, it is possible to deduce, from the measurement of two different components of the electric field, the anisotropy of the wave number spectrum of the turbulence. The method used for this determination will be described and applied to physical situations observed in the polar cap, in the auroral zone, and in sub-auroral regions.

## DISCUSSION

**K.Rawer**

(1) La résolution de vos observations est-elle "satisfaisante" dans le sens où une amélioration technique vers des fréquences plus élevées puisse encore influencer considérablement les résultats obtenus, ou mener à des phenomènes qui échappent à votre expérience?

(2) Quels sont vos résultats obtenus aux latitudes tempérées?

**Author's Reply**

(1) We have systematically stopped our frequency analysis at 200 Hertz, because at higher frequencies we cannot separate the electric field of the electromagnetic waves from the electrostatic wave of the gradient drift turbulence.

(2) At mid-altitudes we have a very small turbulence which is much smaller than at high latitudes. The turbulence is greatest in the auroral zone and at night.

# AURORAL AND SUB AURORAL F-LAYER IRREGULARITY STUDIES IN THE NORTHERN AND SOUTHERN HEMISPHERE DURING THE EQUINOX TRANSITION STUDY SEPTEMBER 16-26, 1984

J. Aarons, Boston University, Boston, MA 02215, USA
J. C. Foster, M.I.T. Haystack Observatory, Westford, MA, USA
L. Kersley, University College of Wales, Aberystwyth, UK
A. S. Rodger, British Antarctic Survey, Cambridge UK

## SUMMARY

A preliminary model detailing the auroral and sub-auroral nightly behavior of F-layer irregularities during the injection and recovery phases of magnetic storms is outlined. Initially the irregularities spread equatorwards from the auroral region. On subsequent nights in those storms where Dst slowly returns to its zero level there is activity at auroral latitudes at the end of the magnetic storm and finally, relatively high irregularity intensity primarily at sub auroral latitudes. The model was tested with observations of the irregularities observed during the September 1984 Equinox Transition Study. Observations of irregularities and incoherent scatter have been examined for the injection phase of the magnetic storm of Sep 22-23. The concept was to determine what parameters are most important in indicating whether irregularities are present over a range of latitudes from 53° CGL to 78° CGL. The data set indicates distinct correlation with convection velocities in the early portion of the storm when a sub-storm was shown on magnetograms.

In the latter half of the night with the magnetic storm fully developed, the data show good correlation with an electron density trough, high electron temperatures, and high velocities. For the recovery phase of a magnetic storm another period in the ETS is examined with sub-auroral irregularities observed during the recovery of Dst when quiet magnetic conditions prevailed. The general form of the model appears to be valid but many aspects of the timing and the magnitude of the irregularity intensity are to be determined.

## I. INTRODUCTION

Irregularities in the F layer at altitudes ranging from 220-550 km are present at sub auroral and auroral latitudes during various phases of magnetic storms. A model has been developed with many features still not validated by observations. The model is designed to show the behavior of F-layer irregularities during the injection and recovery phases of a magnetic storm. The data base for the presence of irregularities is composed of spread-F and scintillation observations. The data base needed to determine which instability processes are operative under various ionospheric conditions and at various latitudes are ionospheric characteristics. These include electron density, electron and ion temperature and convection velocity. These parameters are determined from incoherent scatter radars. The concept in this note is to outline a model of irregularity occurrence at sub auroral and auroral latitudes during and after a magnetic storm. We use case studies from the period Sep. 15-26, the Equinox Transition Study. From these case studies we would expect to determine which of the ionospheric parameters appear to be most closely correlated with the observations of irregularities in the various phases of magnetic storms; however further studies would be needed to test the hypotheses developed

## IIa. MODEL DEVELOPMENT

From several case studies (Aarons et al, 1988) the temporal development of F-layer auroral and sub auroral irregularity activity has emerged. It is simplified and illustrated in Figure 1. The basic concept is that during the early stage of a magnetic storm, auroral processes expand to sub-auroral latitudes producing conditions for creating F-layer irregularities. Lower irregularity intensities are observed at the lower latitudes. On the day of the magnetic storm the F-layer irregularity intensity spreads equatorwards encompassing auroral latitudes plus what were earlier sub-auroral regions; there may or may not at this time be a "trough" in irregularity intensity (Huominer et al, 1981). When the magnetic storm wanes, the recovery phase of the storm occurs. This produces F-layer irregularities of moderate intensity at auroral and sub-auroral latitudes on the night after the main phase of the magnetic storm. When the magnetic activity has completely subsided there may be (as illustrated in the night of Day 2) irregularity intensities of relatively high level at sub-auroral latitudes with little activity at auroral latitudes.

The Dst index exhibits a large negative excursion in the injection phase of moderate and great magnetic storms as shown in Figure 1 which is derived from curves for great and moderate magnetic storms. For the nighttime activity in the recovery phase we have hypothesized that the ring current storage of ions and release of energy as shown by the slow recovery of Dst will create conditions for producing sub-auroral irregularities even when the magnetic index has returned to low values. This has been shown in Aarons, (1987) and Aarons et al (1988) for storms in March, September and October 1981.

It should be noted that irregularities during the day at sub auroral latitudes are of lower levels even during magnetic storms.

### b. THE DATA BASE

The data base for determining the occurrence of irregularities are observations by ionosondes as spread F and by recording scintillations on satellite radio beacon signals. A large number of data sets is available for this study. The map for Figure 2 is an overall picture of the overhead location of the ionosondes and of the ionospheric intersections at 420 km of the satellite beams for stations in the region near 70° W longitude. The conjugate position of the ionosonde at Argentine Islands is also given. The figure includes the probing region of the Millstone Hill Incoherent Radar during the case studies to be considered.

The data sets examined include the following intersections or propagation paths.

| | LATITUDE | LONGITUDE | CGL |
|---|---|---|---|
| Goose Bay, Labrador (Beacon Satellite) | 59° N | 66° W | 70° |
| Halley Bay, Antarctica | 76° S | 27° W | 61° |
| Sagamore Hill, MA USA (Beacon Satellite) | 50° N | 77° W | 61° |
| Goose Bay, Labrador (FLT W) | 48° N | 68° W | 61° |
| Goose Bay, Labrador (GOES 3) | 47° N | 66° W | 60° |
| S. Uist, UK | 57° N | 7° W | 58° |
| Sagamore Hill, MA USA (GOES 3) | 39° N | 74° W | 53° |
| Sagamore Hill, MA USA (FLT W) | 39° N | 74° W | 53° |
| Slough, UK | 52° N | .5° W | 51° |
| Argentine Islands, Antarctica | 65° S | 64° W | 49° |

In addition the scintillation of NNSS satellite beacon signals at 150 MHz and 400 MHz were recorded in Kiruna, Sweden (68° N and 27° E). The satellites' high inclination orbit allowed measurements from 60° to 80° CGL in the sector viewed from Kiruna.

Thus both the latitude and longitude coverages were extensive with observations from 49° CGL to 80° in latitude and 15° E to 80° W in longitude.

## III. THE MAGNETIC STORM

### a. THE MAGNETIC STORM DEVELOPMENT

The magnetic storm of interest took place on the night of Sep. 22-23, 1984. Since most of the observations of irregularities were done in the 70° West region, we shall examine in detail the effects of magnetic activity using magnetograms from Ottawa, Canada, Fredericksburg, Virginia. USA, and Halley Bay, Antarctica. Kp during this period (1200 Sep 22 to 1200 Sep 23) was:

3+ 2- 5+ 5- 5+ 7- 7- 7-

The H component of the Frericksburg magnetogram for the night of Sep 22-23 is reproduced in Figure 3. A sub-storm can be noted starting approximately at 2000-2015 UT (1515 Local Standard Time at 75° W). The AE index peaked between 2030 and 2100 with the upper trace, AU, derived from observatories between 46° W and 78° West. The lower trace was taken from observatories between 80° and 130° East. Strong magnetic activity however appeared to start at 0215 UT with very active variations from 0430 through 1200. We can isolate the two periods in our analysis of 70° West data, characterizing them as afternoon and early evening and midnight-post-midnight. A more dramatic illustration of the sub-storm will appear in Figure 5, a tracing of the magnetogram taken at Halley Bay.

### b. THE AFTERNOON AND EARLY EVENING PERIOD (2000-2230 UT)

In Figure 4 we have plotted irregularity intensity for various sites near 70° W in an effort to determine the timing of irregularity occurrence as a function of latitude. Intensity of irregularities has not been normalized for each of the observations. Ionosonde data are indicated as excursions while scintillation data are indicated in dB or excursion. GOES 3 observations are made at 137 MHz and the other scintillation studies are done at frequencies near 250 MHz.

1. The Higher Latitudes: In the 70° West region irregularities appeared at the higher latitudes at 2015 UT. This included observations from Goose Bay for two paths at 70-75° West. For the Halley Bay longitude, strong spread F began at 1830 UT and probably continued but was obscured by auroral absorption (Rodger et al 1981).

2. The Lower Latitudes: Using Figure 4 and concentrating on 70° W data the start of scintillation activity was 2015 UT for 60.5° (GOES GB) and 2130 for 53° (GOES SH) and 2300 for 49° (Argentine Islands). There was an equatorward time displacement. The timing of the appearance of a moderate intensity level in the 70° W region is given in Figure 5 along with a portion of the Halley Bay magnetogram. Throughout the paper we use conjugate points interchangeably to determine general activity patterns of both the magnetic field and the F-layer irregularities (Rodger and Aarons, 1988). The very rough velocities of the initiating mechanisms yielded an equatorward velocity of 250 m/sec in the region from 60.5 to 53° and 100 m/sec in the region from 53° to 49°.

3. However for two paths there were low irregularity intensities. MARISAT recordings at 52° W from Goose Bay showed little displacement in this time period. This is indicative of sub storm behavior with irregularities localized in narrow longitude sectors. For the FLTSATCOM (E) recordings from the Boston area at 63° W, but at low latitudes no scintillations were observed in this time period. However, the observations at 250 MHz are of considerably lower sensitivity than those from 136 MHz (GOES 3) and probably from the ionosonde's recordings of the presence of spread F. Thus for this observation the high frequency of the satellite beacon, the confinement of the irregularity region, and the propagation path's displacement may be the reasons for the low level of irregularity intensity on this path. In addition Slough failed to note spread F during this early time period. There are indications that the irregularity intensity extended farther equatorward in one longitude region than in the region from 0-60° West. S. Uist, polewards of Slough, showed spread F throughout this time period.

### c. KIRUNA OBSERVATIONS

The start of the operational program to observe satellite scintillations took place on this date for the entire program. Thus the data are not available before 2100 UT. A contour map of the results of observing 150 MHz signals from the NNSS satellites is shown in Figure 6a. Probably the lowest latitude of the .2 contour of S4 occurred at 61° near 2100. The higher contours indicate that high levels of irregularity intensity probably occurred after the start of observations. Both the .2 and .4 contours move poleward between 23 UT and 0130 UT. The local time difference between the Kiruna data set and the 70° West meridian is approximate 6-7 hours yet the descent of the irregularity regions occurs at the same UT as the 70° West measurements. This is indicative of the effect of magnetic activity on lowering the latitude of the region where irregularities are observed and on increasing the intensity of the F-layer irregularities. In examining Lerwick magnetograms the substorm and beginning of the magnetic storm are quite similar to those from Frericksburg and Ottawa. The response is clearly due to the sub-storm activity.

d. THE LATER TIME PERIOD

After a period of relative quiet on the magnetograms, the magnetic field became more variable at 0215 UT on the Ottawa and Fredericksburg magnetograms. Extremely strong variability was observed from 0445-0630 and from 08-09VT. The generally high level of irregularity intensity at auroral latitudes can be readily seen in Figure 6b, the contours of irregularity intensity at 150 MHz as observed from Kiruna.

The Northern latitude group showed scintillation activity 0530-1000. At the beginning of magnetic activity the irregularities intensities became apparent.

Increased irregularity intensities were noted at the lower latitudes in the 70° West sector beginning at 0245 (Slough) and particularly during 05-12 UT depending on longitude and latitude. The Argentine Islands ionosonde showed intense irregularities 0630-10 UT. The effect of increased magnetic activity is apparent.

## IV. INCOHERENT SCATTER RADAR DATA

The occurrence of F-layer irregularities as a function of latitude near the 70° W meridian can be related to the state of and changes in the local ionosphere by comparing irregularities with the F-region observations made by the Millstone Hill incoherent scatter radar. In this Section we present latitude maps of convection velocity vectors (Figure 7) and of ionospheric temperature and density (Figure 8a) for the 24-hour interval beginning at 1600 UT on Sep 22, 1984. The occurrence of scintillations and spread F during this interval is indicated in Figure 8b. Before discussing the details of this combined data set several aspects concerning the correlation of the radar-deduced parameters with the observed levels of F-layer irregularity intensity should be considered. One is the displacement of the intersections of the 420 km ionospheric intersection point in longitude with respect to the probing area of the radar. This will lead to discrepancies in timing of phenomena. Secondly the time for instabilities to fully develop under a set of circumstances is difficult to assess until the precise mechanism is understood. Finally as pointed out by Kelley et al (1982) and Vickrey et al (1982) the duration of the F-layer irregularity regions is lengthy so that persistence plays a role in establishing a correlation of phenomena. The attempt in this note is to point out similarities in development of irregularities and F-layer characteristics.

The observing mode used at Millstone Hill during the ETS experiment consisted of paired elevation scans with the 150-foot steerable antenna from north to south along adjacent meridians. This observing technique has been described in detail by Oliver and Salah (1988). Sets of line-of-sight velocity observations at roughly the same magnetic latitude were combined to determine the vector plasma convection velocity at each latitude with roughly 45 minute time resolution. These data are presented in Figure 7 in clock-dial geodetic latitude-magnetic local time format and serve to indicate the direction of the plasma convection. Contour plots of the magnitude of the convection velocity are presented in Figure 8a for a more direct comparison with the irregularity occurrence data shown in Figure 8b. That figure is an overlay of scintillation and spread-F activity with intense irregularities shown as solid fill and dotted fill indicating moderate levels for a particular frequency and observatory.

1. Convection

The earlier period of strong local irregularities (21-24 UT on Sep 22) coincided with a strong enhancement of the afternoon sector westward convection velocity which extended equatorward to near 40° latitude. There is no indication of higher latitude irregularities being convected equatorward. The relatively high convection velocities at 46-48° geographic latitude, the GOES 3 intersection from Goose Bay, can also be seen in the contours beginning at or just before 2000 UT the start of the substorm (Figure 8a). It should be noted that the propagation angle intersection latitude is a function of the altitude chosen for the oblique path. At 2200 there are isolated velocities of 150 m/sec reported for the latitudes equatorwards of Millstone Hill. As shown in Figure 4 the 60°-62° CGL high latitude scintillation starts near 2030 while the lower latitude scintillation and spread F commence from 2230 to 2300, the observed time of the short period of the lower latitude high convection.

A second interval of enhanced convection velocity begins at the highest latitudes sampled near 24 UT and progresses equatorward reaching its lowest latitudes between 0030 and 0200 UT. At all latitudes the created irregularities endured during this time period even when the convection diminished. Velocities decreased between 0230 and 0500 UT at all latitudes as the convection reversal was crossed (although some velocity data are missing within the trough (cf Figure 8e) where the density was very low). For 53° scintillations stopped at 02 UT; the Argentine Islands intersection did not observe spread F for several hours after 24 UT. These observations suggest that if the convection velocity is involved in generating the irregularities, there is a slight delay in creating the conditions responsible for the irregularities. The latitudinal extent of the enhanced velocities is limited by the shielding of the inner magnetosphere, but time dependent effects near the inner edge of the ring current (eg Foster and Aarons, 1988) could be responsible for velocity-dependent irregularity formation at relatively low latitudes during dynamic events.

Strong eastward velocities were observed poleward of Millstone Hill after 06 UT while erratic measurements were seen in the low density trough region to the south. A second period of strong irregularities was associated with the enhanced convection velocity and was seen at the lower latitude observing points as the enhanced velocity progressed equatorward along the meridian. In this part of the night Slough showed spread F from 0245-0500 UT, possibly indicating that the convection velocities were very high at that longitude during the interval. After 1030 UT convection velocities observed from Millstone Hill decreased significantly as did the irregularity intensity observed along the 70° KW meridian.

2. Electron Temperature

Contours of electron temperature in the topside F region along the Millstone Hill meridian are shown in Figure 8c. A data gap is collocated with the trough and extends from 52° at 01 UT to 38° at 05 UT. Although the electron temperature increases dramatically as the density drops as the trough is entered at a given position, no specific features of the temperature data appear to correlate with the occurrence of irregularities shown in Figure 8b.

3. Electron Density

A deep ionospheric total density trough progressed equatorward along the Millstone Hill meridian during this event as is often the case during disturbed evening sector conditions. Figure 8d presents iso-density contours which describe the temporal variation of the density/altitude profile overhead at Millstone Hill (42° latitude) through the event. The trough passed over the station at 0400 UT and represented a decrease in density of at least a factor of 10 at each altitude. Figure 8e presents contours of the F-region peak density in the format of Figure 8b which clearly show the equatorward progression of the trough along the 70° W meridian. The trough remained 3°-4° wide and the F region was enhanced by a factor of 10 on its poleward side. The contours of F-region density indicate that during the first appearance of lower latitude F-region irregularities (21-02 UT) the stations making the observations were all equatorward of the trough and there were no large-scale density blobs apparent. It is noteworthy that the minimum in irregularity amplitude observed at each station (Figure 8b) between 00-06 UT occurred as the density trough passed over each site. The second enhancement of irregularities occurred as the stations sampled the ionosphere poleward of the trough.

## V. CONTRAST IN SUB-AURORAL IRREGULARITIES DURING THE RECOVERY PERIODS OF A MAGNETIC STORM AND DURING AN EXTENDED MAGNETICALLY QUIET PERIOD

An example of the difference in sub-auroral F-region irregularity intensity in the recovery period of a magnetic storm (Sep 21-22) compared with the extended quiet period of Sep 15-19 could be found in this data set.

The comparison was made in the following manner. The maximum excursion of scintillation, the maximum scintillation index, or the maximum frequency spread as noted on the ionograms that was recorded for each individual observatory during the entire time period was given the value of 1. This was in most cases the value reached during the magnetic storm of Sep 22-24. The ratio of irregularity intensity in the midnight time period of each of the three nights to this high level is plotted in Figure 9. The nights chosen represented the various stages of a magnetic storm; thus we have plotted the relative change of irregularity intensity. The period encompassed:

(a) a night (Sep 18-19) when there had been a long period of magnetic quiet (Kp for the night was 2) when only low level irregularity intensity was observed across the network of stations

(b) a storm day (Kp=4+) (Sep 19-20) when irregularity intensity was high

(c) a night (Sep 21-22) when Kp was low (2) but the sub-auroral region had relatively high levels of irregularity intensity. Thus considerably more intense irregularities are observed during storm recovery compared with a day of similar geomagnetic activity (Sep 18-19) prior to a storm.

The data showed that the night of Sep 21-22 with low magnetic activity in the recovery phase of the magnetic storm had levels of irregularities in the sub-auroral region similar to those of a magnetically disturbed night (Sep 19-20). A night before the onset of magnetic activity (Sep 17-18) with the same level of magnetic activity as the recovery night had little or no irregularity intensity at sub-auroral latitudes.

The month of September 1984 was a period with several magnetic storms and large displacements of Dst even though the solar flux was relatively low. The general activity pattern can be seen in Figure 10 by the plot of scintillation levels at 53° CGL from the GOES satellites. The agreement between the pattern of Dst levels and sub-auroral intensity levels can readily be seen from this diagram. The active periods of the months are clearly distinguishable from the quiet periods when Dst was low. In line with the model one can note that sub-auroral irregularities exist not only during periods when Dst is actively going negative but during the slow recovery of Dst to its zero value. In addition however one can note that the model is highly idealised.

## VI. DISCUSSION

The intensity and occurrence of F-region irregularities at auroral latitudes has been covered by earlier work (Rino and Matthews, 1980, Aarons, 1982). Studies using scintillations as recorded at Slough (Bramley, 1974) have shown the lack of correlation with magnetic indices. In-situ studies have shown the presence of sub-auroral irregularities (Clark and Raitt, 1976).

The model presented has different phases of a magnetic storm acting probably under different instability conditions. Just as in pictures of the development of predominantly 100 km auroral forms, during the initial phase of a storm there is an expansion equatorwards (we do not deal with the expansion polewards). The F-layer irregularities gradually encompass what were sub-auroral latitudes with levels lower than the intensity of the auroral irregularities.

We have shown (Aarons et al, 1988) that periods of relative high irregularity intensity normally occur at sub-auroral latitudes during the recovery phase of a geomagnetic storm when Dst is appreciably negative; irregularities are rare when Dst hovers near 0 nT. The ring current is the major source of free energy in the magnetosphere and plasmasphere that could provide the necessary additional energy for plasma instabilities to occur in the sub-auroral F layer. There is considerable complementary evidence for this conclusion. For example, stable auroral red arcs (SAR arcs) are a common occurrence for several nights after significant geomagnetic disturbances. In addition it has been shown that F-layer irregularities are closely associated with SAR arcs (Rodger, 1984; Aarons, 1987 and the references therein). The energy for the SAR arcs comes from the ring current (Kozyra et al, 1987, Slater et al, 1987). Thus the ring current acts as an energy reservoir for several days after geomagnetic activity (Kp and Ap) has diminished. It is proposed that the ring current energy provides the source for creating the instabilities at sub-auroral latitudes as well as SAR arcs.

We do not propose a simple correlation between Dst and irregularity intensity at sub-auroral latitudes. The ring current consists of protons, oxygen, helium and other ions with injection for any storm differing both from the percentage of added ions and their energies. In addition both charge exchange and ion cyclotron waves produced by scattering of ring current ions have been proposed as possible mechanisms for the gradual return of the ring current to its quiet level. Until a definitive mechanism has been accepted for the generation of the sub-auroral irregularities, one cannot use simple correlation studies.

As for the correlation of high convection velocities and the occurrence of irregularities one example is available of primarily daytime observations (Foster and Aarons, 1988). The generation of irregularities during shear conditions has already been shown (Basu et al, 1984) and it appears likely that the high convection levels produce the necessary conditions during these storm periods.

In a review (Tsunoda, 1988) found that the very large scale structures in the polar ionosphere were produced by magnetic-flux-tube interchange processes with particle precipitation and some kind of wave activity contributing to the smaller scale. However the strong irregularities, Tsunoda concluded, appear to be controlled by the E x B instability. Auroral irregularities are produced near auroral boundaries where spatially varying convection allows significant slip velocities to develop and where the E region Pedersen conductance is low. Irregularities also may be generated under conditions when the electric field is time-varying with periods less than an hour (substorms and pulsation events) provided that the underlying E layer is not highly conducting.

## SUMMARY

This set of data indicates that F layer irregularities in their time of appearance and in the latitudinal extent are correlated with high convection velocities recorded at their latitudes. In the early evening portion of the data set, there was no indication of irregularities arising from auroral activity and moving equatorwards nor was there in the midnight and post-midnight periods since the convection pattern north of MH was mostly EW. However there were consistently high EW velocities at the lower latitudes when scintillations or spread F was observed. A slight increase in electron density was noted at 275 km during the early period of scintillations (at 42° Latitude) but the increase scarcely resembled the "blob" as depicted in the literature.

## ACKNOWLEDGEMENTS:

Support for this study for J. Aarons has come principally from the Office of Naval Research (US) with partial support from the National Science Foundation (US). Millstone Hill Radar operations and analysis were supported by the National Science Foundation through Cooperative Agreement ATM-8419117 to the Massachusetts Institute of Technology.

## REFERENCES

Aarons,J. (1982) Global morphology of ionospheric scintillations, Proc. I.E.E.E., 70, 360

Aarons,J. (1987) F-layer irregularity observations of the SAR arc event of March 5-6 1981, Radio Sci., 22, 100, 1987.

Aarons, J. C. Gurgiolo, and A.S. Rodger (1988) The effects of magnetic storm phases on F-layer irregularities below the auroral oval. Accepted for publication in Radio Science

Basu, Su., S. Basu, E. MacKenzie, W. R. Coley, W. R. Hanson, and C. S. Lin (1984) F region electron density irregularity spectra near auroral acceleration and shear regions, J. Geophys. Res. 89, 5554

Bramley, E.N., (1974) Fluctuations in direction and amplitude of 136 MHz signals from a geostationary satellite J. of Atmos. and Terr. Phys. 36, 1503, 1974.

Clark, D.H. and W.J. Raitt (1976) The global morphology of irregularities in the topside ionosphere, as measured by the total ion current probe on ESRO-4, J. Planet. Sp. Sci. 24, 873

Foster, J.C., C. del Pozo, K. Groves, and J.-P.St. Maurice, (1988) Radar observations of the onset of current driven instabilities in the topside ionosphere, Geophys. Res. Lett., 15, 160

Foster, J.C. and J. Aarons (1988) Enhanced anti-sunward convection and F-region scintillations at mid-latitudes during storm onset J. Geophys. Res. Accepted for publication

Houminer, Z., J. Aarons, and F. Rich, (1981) Production and dynamics of high-latitude irregularities during magnetic storms, J. Geophys. Res. 86, 9939

Kelley, M.C., J.F. Vickrey, C.W. Carlson and R. Torbet, (1982) On the origin and spatial extent of high-latitude F-region irregularities, J. Geophys. Res. 87, 4469, 1982.

Kozyra,J.U., E.G. Shelley, R.H. Comfort, L.H.Brace, T.E. Cravens, and A.F. Nagy (1987) The role of ring current O+ in the formation of stable auroral red arcs, J Geophys. Res., 92, 7487

Rodger, A.S. and J. Aarons (1988) Studies of ionospheric F-region irregularities from geomagnetic mid-latitude conjugate regions J. Atmos. and Terr. Phys. 50, 83

Slater,D.W., C.Gurgiolo, J.U.Kozyra, E.W.Kleckner and J.D.Winningham (1987) A possible energy source to power stable auroral red arcs: precipitating electrons, J. Geophys. Res., 92, 4543

Tsunoda, R.T. (1988) High-Latitude F-Region irregularities; a review and synthesis to be published in Physica Scripta

Vickrey, J.F. and M.C. Kelley (1982), The effects of a conducting E layer on the classical F-region, cross-field plasma diffusion, J. Geophys. Res. 87, 4461

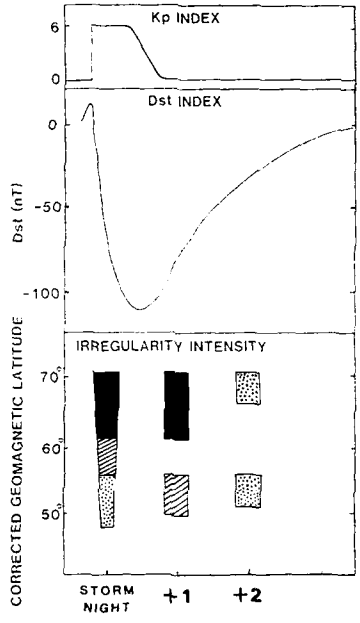
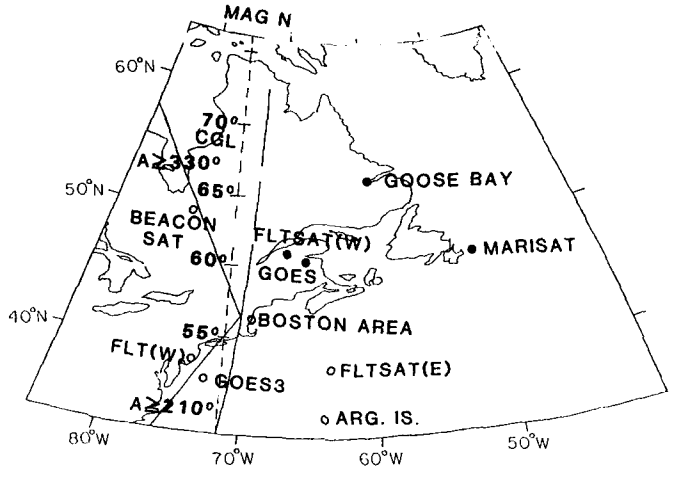



Figure 1. Picturization of F-layer irregularity intensity at auroral and sub-auroral latitudes during a magnetic storm. Initially the auroral latitudes show intense irregularities at the beginning of the storm. With a time delay the F-layer irregularities expand but the intensity diminishes away from the auroral region. With the subsiding of the magnetic activity the auroral irregularities decrease in intensity, the sub-auroral irregularities exist even with lower magnetic activity. Dst slowly returns to normal; sub-auroral irregularities slowly return to lower levels.




SEP 84

Figure 2. Map of some of the ionospheric intersections (at 420 km for the satellite beacons) and the ionosondes used in this paper. Coverage area of Millstone Hill incoherent scatter observations for this period is also given. Millstone Hill took data at 330°, 0°, 180° and 210° azimuths. The scintillation observations from the Boston area are open circles. The scintillation observations from Goose Bay are solid dots.

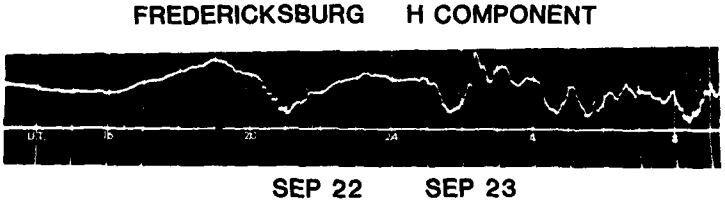



Figure 3. The H component of the Frericksburg magnetogram for the period of interest Sep 22-23, 1984.

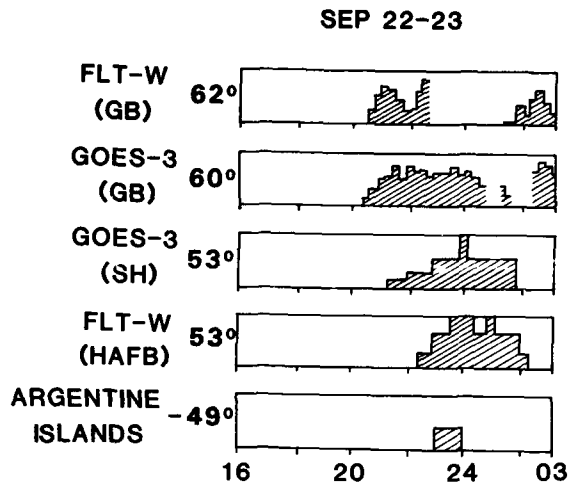



Figure 4. Spread F and scintillation data for the 70° West longitude region for the period of the magnetic bay (2000-2300 UT). See also Figure 5. Amplitudes shown are only relative to maximum values observed for a particular location and a particular observing frequency. The graph is developed to illustrate the temporal development of activity rather than to compare absolute values of the irregularity intensity.

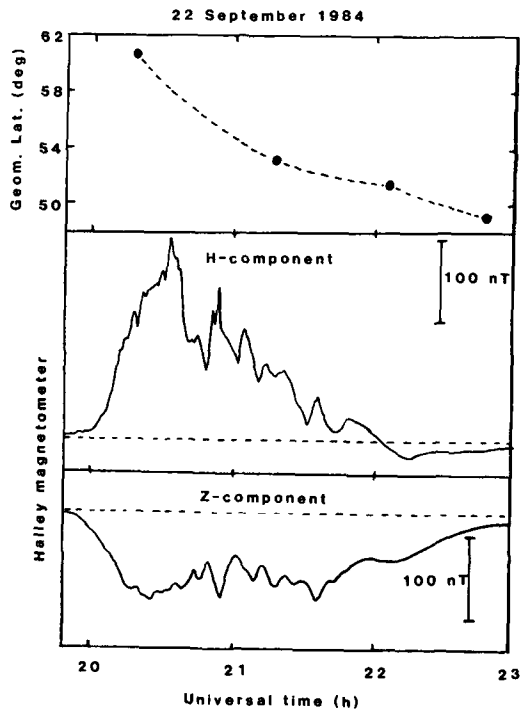
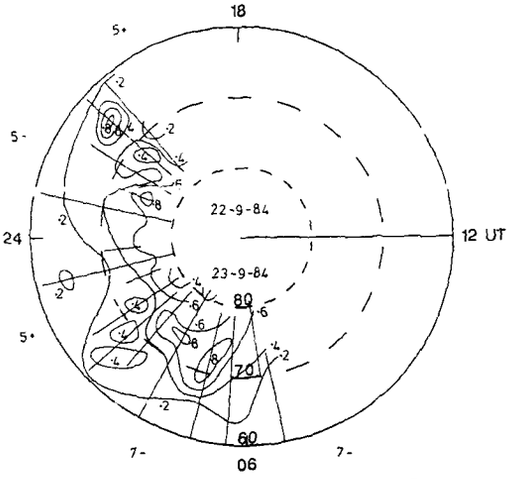



Figure 5. Time displacement of irregularity occurrence at 2100 UT Sep. 22-23 in the 70° W sector. The equatorward delay in the start of the irregularities is graphed. The Halley Bay magnetogram for the period is also shown.

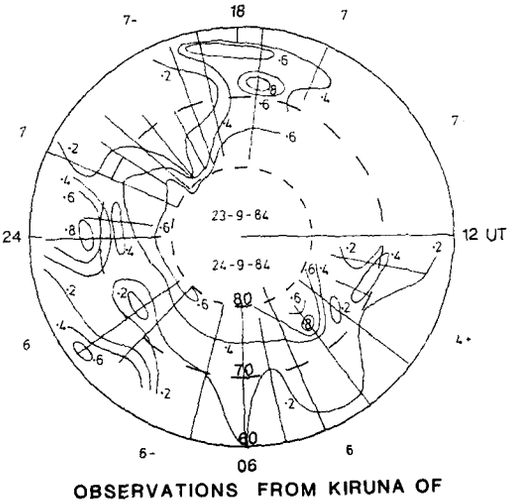



OBSERVATIONS FROM KIRUNA OF

150 MHz SCINTILLATIONS (S4)

SEP 22-23, 1984

Figure 6a. Contour map of scintillation indices (S4) at 150 MHz with data taken at Kiruna, Sweden. Data is for the night of Sep. 22-23




OBSERVATIONS FROM KIRUNA OF

150 MHz SCINTILLATIONS (S4)

SEP 23-24, 1984

Figure 6b. Scintillation levels for data taken at Kiruna on the night of Sep. 23-24.

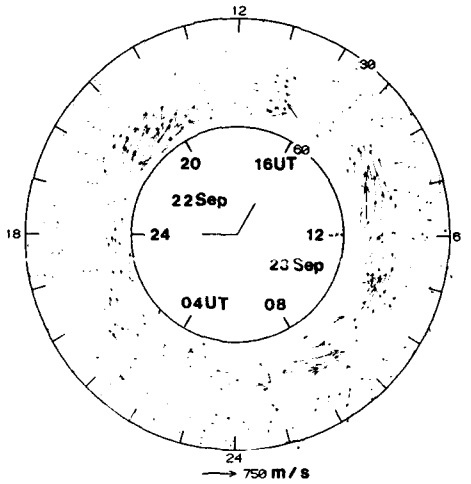



Figure 7. Millstone Hill convection diagram for the night of Sep 22-23 The time period encompasses 1600 UT on Sep 22 to 1600 UT on Sep 23. Local time is noted on the outer edge of the diagram. Geodetic latitudes are plotted.

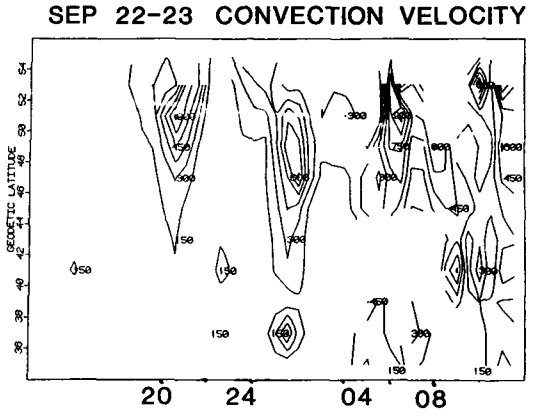



Figure 8a. Convection velocities as observed by the Millstone Hill radar as a function of latitude.

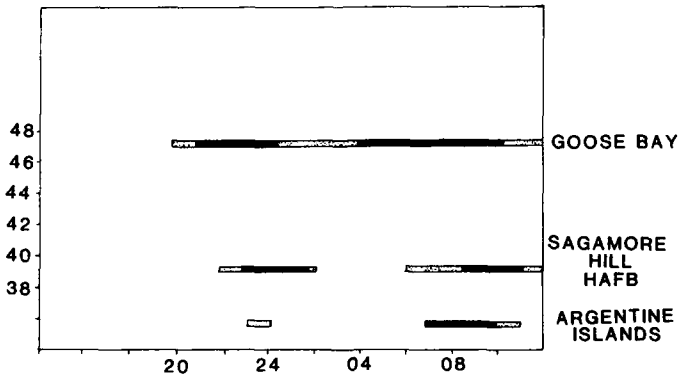



Figure 8b. Scintillation and spread F activity during the early portion of the Sep 22-23 night.

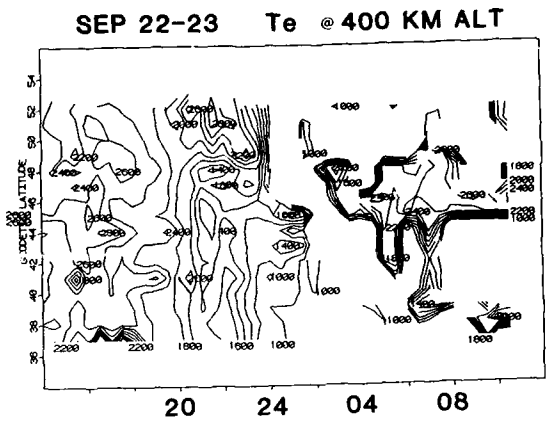



Figure 8c. Te contours for the period.

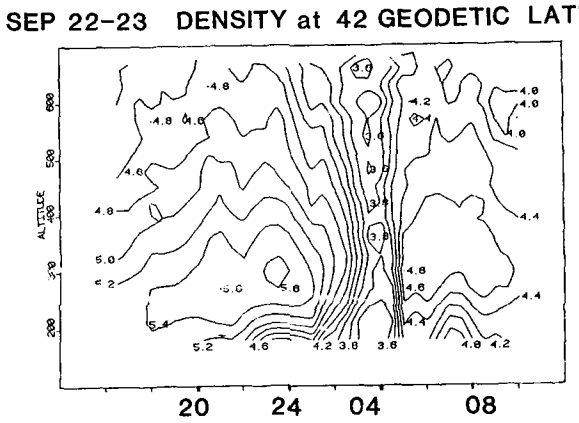



Figure 8d. Electron density at 42° Geodetic Latitude

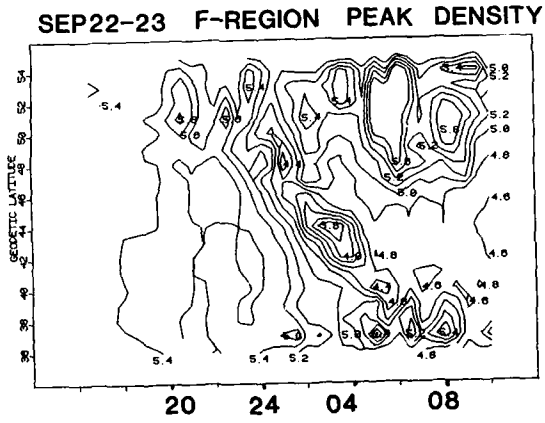



Figure 8e. Peak electron density

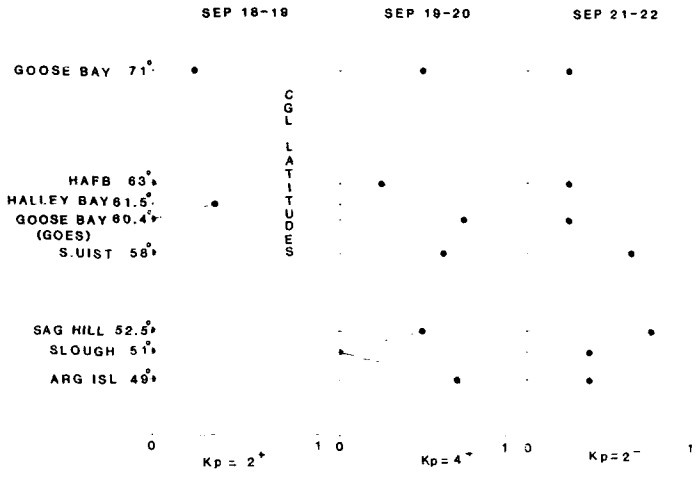



Figure 9. The relative occurrence of irregularities over 20° for three local midnight periods in September 1984. The data were normalized for each propagation path separately with a value of 1 as the maximum level observed over the period Sep 16-26, 1984. The nights are representative of a quiet period (18-19), a geomagnetically active period (19-20), and a storm recovery period (21-22). The relatively high levels of irregularity intensity can be noted for the night of 21-22 compared to the night which had similar magnetic activity (18-19).

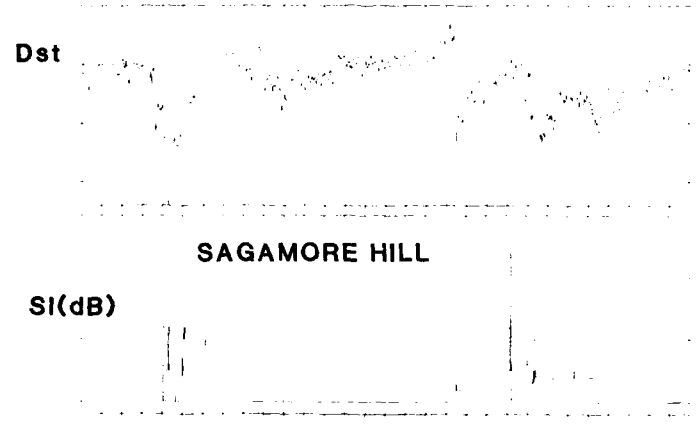



Figure 10. Dst and Scintillation Index variations (Sagamore Hill GOES 3 at 137 MHz) for the month of September 1984

## DISCUSSION

**P.Vila**

Do you think maintenance of strong irregularities has a LT dependence? This could be important at mid-latitudes where the $E$ layer conductivity might cut off the electric fields necessary to maintain $F$ region irregularities.

**Author's Reply**

We think so, and are looking into this dependence, which is so strong at equatorial latitudes.

# SMALL-SCALE IRREGULARITIES AND GLOBAL PLASMA DYNAMICS IN THE HIGH-LATITUDE IONOSPHERE

**C.HANUISE**
L.S.E.E.T./C.N.R.S.
Université de Toulon et du Var
639, Bld des Armaris, 83100 Toulon, France

**R.A. GREENWALD, K.B. BAKER, J.M. RUOHONIEMI**
Applied Physics Laboratory
Johns Hopkins Rd, Laurel, MD 20707, USA

**J.C. CERISIER, C. SENIOR**
Centre de Recherches en Physique de l'Environnement
4, Av. de Neptune, 94100 St. Maur des Fossés, France

## SUMMARY

New sophisticated HF coherent radars have been installed in the past few years at very high latitudes. The APL Goose Bay radar and the French SHERPA radar in Schefferville are the two components of the Polar Region Ionospheric Structure Monitor (PRISM) operating in Canada. These instruments detect irregularities within the three-dimensional volume covering much of northeastern Canada and Greenland. They are capable of following the temporal variability of these irregularities as well as determining unambiguously the Doppler shift and spectral broadening of radar signals scattered by them.

The scientific goals of the program concern both the physics of the small-scale structures and the mapping of large-scale plasma convection. Several case studies of plasma mechanisms generating the small-scale density structures have already been published. The temporal and spatial variability of the large-scale irregularity structures, as seen by the radars, is also investigated. The relationship between the radar Doppler velocity and the plasma drift velocity has been confirmed.

We emphasize in this paper the potential of the system for studying global ionospheric dynamics. When combining the two radial velocities measured by each radar, one derives a two-dimensional map of plasma convection and electric field over a region as large as 2 million square kilometers, with a fine temporal and spatial resolution in the auroral region and over the polar cap.

## INTRODUCTION

The coherent radar technique is now well established as a tool for studying both plasma structure and ionospheric dynamics of the high-latitude ionosphere. In the VHF frequency band, the STARE system has pioneered observations of E region phenomena (Greenwald et al., 1978) and similar instruments (SABRE, BARS) have been developed in other parts of the world. Simultaneously, new sophisticated radars, operating at HF frequencies, have been installed in the auroral and polar regions (Hanuise et al., 1981; Greenwald et al., 1983, 1985). They complement the instruments operating at higher frequencies by extending their latitude and altitude coverage.

It is indeed well known that coherent radars detect backscatter from ionospheric irregularities which are aligned along the earth's magnetic field. As a matter of fact, they perform a spatial Fourier transform of the density fluctuations present at various scale sizes and pick up those with a wavelength equal to half the radar wavelength. Moreover, the detection is possible in backscatter only when the incident wave vector becomes normal to the magnetic field. This condition, called aspect sensitivity or normality condition, can be reached more easily with HF radarsdue to the wave refraction (Fig. 1).

At high latitudes, the earth's magnetic field is nearly vertical and its dip angle increases with increasing latitude. As a consequence of the earth's curvature, the normality condition can only be achieved if the ray path is refracted towards the horizontal as the wave enters the ionosphere . The wave transmitted by a VHF radar has a frequency well above the maximum local plasma frequency, and therefore propagates along a straight line. It does not fulfill the aspect sensitivity condition, except in the auroral E region, where all VHF radars operate. HF radars do not suffer the same limitation, as their operating frequency is of the order of the local plasma frequency. The refraction is then sufficiently large that the wave reaches perpendicularity to the magnetic field. Such radars operate therefore successfully in both the E and F regions at auroral and polar latitudes.

Observations of F-region irregularities with modern coherent HF radars were reported first with the SAFARI (Scandinavian And French Auroral Radar Investigation) system in Scandinavia (Hanuise et al., 1981). Similar observations were performed by the Applied Physics Laboratory group in Alaska, using a computer controlled system (Greenwald et al., 1983). As a follow on, a cooperative program was set up around two similar HF radars installed in northern Canada (Greenwald et al., 1985). This program, known as PRISM (Polar Region Ionospheric Structure Monitor) is aimed at studying both the physics of the ionospheric plasma and the electrodynamics of the very high latitude ionosphere. In this paper, we study the derivation of global scale patterns of plasma convection from preliminary data obtained with the PRISM system. Firstly, we present the technical characteristics of the two radars and the scientific objectives of the program. Then we look into the problems related to the derivation of the bulk plasma drift from the radar Doppler velocities. We also present data for several periods and discuss the validity of the two-dimensional maps of the plasma flow. Finally, we emphasize the potentiality of such radar systems for studying large scale ionospheric dynamics.

## DESCRIPTION OF THE RADAR

The PRISM program consists in two HF coherent radars operated by the Johns Hopkins University/APL and several french groups, respectively in Goose Bay, Labrador, and Schefferville, Quebec. The APL radar has been in operation since october 1983 while the Schefferville radar has been installed later, with limited observations starting in early 1986. Both radars have basically the same characteristics which were detailed in a paper by Greenwald et al. (1985). Therefore, we shall only describe here the main parameters of the system.

The radars are capable of broadband operations between 8 and 20 MHz, such as to optimize the ray propagation and achieve the normality condition. Indeed, the high latitude ionosphere is affected by both diurnal variations and geomagnetic disturbances. The associated variability in the electron density causes an HF wave to undergo different amount of refraction in the course of a day. It may be necessary to operate a radar over an extended frequency range to determine the best frequency for backscatter observations, or to examine the entire ionospheric volume at any particular time. The broadband operation depends on the radar itself, and on the antenna array which is made of 16 log-periodic antennas (8 in Schefferville) used for both transmitting and receiving. The associated azimuthal resolution is of the order of 3.5 degree (7 deg. in Schefferville). The beam formed by the array may be directed into one of 16 possible directions, covering an azimuthal sector of 50 degrees centered on magnetic north in Goose Bay. Such a phasing system is not yet operating in Schefferville, where the beam is directed in the azimuth 15° East. The resulting field of view of the two radars is shown in Fig. 2. One can see that the viewing area covers a very large part of the high-latitude ionosphere and encompasses the nominal field of view of the Sondre Stromfjord incoherent scatter radar.

The two radars are fully computer controlled, for the definition of operating parameters (frequency, viewing direction, integration time, pulse length...) as well as for the digitizing and preprocessing of the data. Transmission is usually performed in a multipulse mode with 7-pulses, giving access to 16 lags of the auto-correlation function (ACF). Individual transmitters are used for each antenna, with a peak power of 1 kW at Goose Bay and 400 W in Schefferville. The ACF values are computed in real time and written on digital magnetic tapes together with the radar parameters. Other parameters of interest are the spatial resolution, usually 20 km in the radial direction and 3.5 deg. in the azimuthal direction, and the temporal resolution usually 5s for observation along one beam and hence 80s to derive a map comprising the 16 available viewing directions. The Goose Bay radar is also capable of measuring the elevation of the incoming backscattered wave through the phase difference between the antenna array and a second array located parallel to the first one and located at a distance of 100 m.

The scientific goals of the program are related to the local physics of ionospheric irregularities and to the global scale electrodynamics of the ionosphere. Indeed, several studies have already been achieved with the Goose Bay radar alone in these fields (e.g. Baker et al., 1986; Ruohoniemi et al., 1987; Villain et al., 1987). Both E and F region irregularities are detected with the PRISM radars. Among the problems to be addressed are their generating mechanisms, the relation of 10-m wavelength irregularities with larger scale structures, the respective role of precipitations and transport in the ionization structure. The irregularity motion is also used to map the plasma convection on a global scale. To be fully explored, the subject needs to use the scanning capability on the two radars. We have nevertheless started preliminary studies in the present configuration, with the Goose Bay radar scanning in 16 directions over 50 degrees in azimuth and the Schefferville radar looking in a fixed direction, intersecting the other radar field of view.

## DERIVATION OF TWO-DIMENSIONAL MAPS

It may seem straightforward to derive two-dimensional maps of plasma velocity from the radial Doppler velocity measured by each radar, but, in the present case, we are only able to combine a limited number of points, located at the intersection between the fixed beam of the Schefferville radar and the field of view of the Goose Bay radar. A specific method should be designed to construct velocity vectors over the whole field of view. Moreover, two questions come to mind when mapping the plasma flow with HF coherent radars. The first one concerns the equality between the Doppler velocity and the ion drift, whereas the second addresses the problem of geographical localization of the backscatter.

Few instruments derive the ionospheric plasma flow from simultaneous observations along two different directions. Considering incoherent scatter radars, only EISCAT has this capability, due to its three receiving stations. Both Millstone Hill and Sondre Stromfjord radars use measurements performed with a single station over a number of azimuths to compute two-dimensional representations of the plasma drift (e.g. Banks and Doupnik, 1975 ; Holt et al., 1984). These techniques have certain limitations that affect the accuracy of the results. The limitations are related to the temporal resolution of the measurements and to the fact that the transverse component of the velocity must be inferred from the azimuthal variation of the line of sight component over a finite spatial volume. Any non-uniformity in the flow pattern may introduce considerable errors in the final result. A similar method has also been used for the Goose Bay data (Hanuise et al., 1985). The initial estimate of the velocity was deduced from the azimuthal variation of the radial component over the field of view. Successive computations used the hypothesis of a divergence-free flow to derive velocity vectors on adjacent beams. The possible effect of the closure of the flow by field aligned currents were therefore neglected. The conclusion was that the validity of the final map was highly dependent on the validity of the initial estimate. In the present case, this initial estimate is obtained by the combination of the radial velocities measured by each radar. Its spatial and temporal resolutions are those mentionned previously for the radar, i.e, typically, 20 km in the radial direction and 80s. The two-dimensional drift analysis will use this accurate determination of the initial value of both transverse and radial velocity components, and we can therefore expect the resulting velocity map to represent accurately the effective convective flow pattern.

Mapping of the plasma flow is possible from the Doppler motion of ionospheric irregularities if they move with the ambient medium. In the E region, the Doppler velocity of irregularities produced

by the two-stream instability is limited to a value near the ion acoustic velocity, which only bears an indirect relation to the plasma drift velocity (Nielsen and Schlegel, 1985). This relationship is certainly true on a statistical basis but might lead to misestimate the drift velocity in specific cases. For this reason, we shall deal here with observations performed in the F region. At these altitudes, the motion of both electrons and ions is due to the magnetic field. Theories, as well as numerical simulations, predict the existence of instabilities with a low phase velocity in the plasma frame of reference (Chaturvedi and Ossakow, 1979 ; Keskinen et al., 1980). The observed Doppler motion is therefore exclusively related to the motion of irregularities with the bulk plasma. A limited nomber of experimental studies have been aimed at verifying this property. Using data collected simultaneously with the SAFARI coherent scatter radars and EISCAT, Villain et al. (1985) compared a limited number of velocity vectors and found strong evidence of the equality of plasma drift with ion drift velocities. In a later experiment, (Ruohoniemi et al., 1987), radial Doppler velocities measured by the Goose Bay HF radar were compared to radial velocities estimates from the Sondre Stromfjord incoherent scatter radar. Over a large number of data comparisons, including velocity reversals exceeding 1000 m/s, an excellent correspondence was found between the measured irregularity and ion drift velocities (Fig. 3). Altogether these results indicate that the small-scale F-region irregularities detected with HF radars can safely serve as tracers of the ionospheric convective drift.

Another question coming to mind when dealing with HF propagation is concerns the localization of the scatter. The refraction of the HF wave, which makes it possible to satisfy the aspect-sensitivity condition, makes it also difficult to derive the geographical location of the radar target from the measured time delay of propagation. VHF/UHF radars are not prone to this uncertainty, as it is straightforward to map time delays into geographical grid with a straight line propagation. Ray tracing techniques have been used extensively to determine the contours of perpendicularity of the radar wave with the earth's magnetic field. They have used simple ionospheric models (e.g. Millman, 1969, 1975 ; Crochet et al., 1971 ; Swenson, 1972), electron density values obtained from ionosonde measurements (Hunsucker, 1969), or meridional profiles of density measured with an incoherent scatter radar (Villain et al., 1984). The irregular structure of the high-latitude ionosphere may induce large deviations in both radial and azimuthal directions and, consequently, makes difficult to localize precisely the scatter. Nevertheless, it is possible, in a first approximation, to use a straight line propagation to map geographically the radar echoes, the shift in range being, in most cases, of the order of a few tens of kilometers. Moreover, the shift will be similar for all ranges, and a large scale map will be globally shifted by this distance.

We have used a simple ray tracing computer program (Jones and Stephenson, 1975) to illustrate these various points. The ionospheric density is modelled by two parabolic layers, respectively at E and F region altitudes. Latitudinal gradients can be included to represent density blobs present in the high-latitude ionosphere. Figure 4 shows the result of this ray tracing when no gradients are included. Parameters of the plot are maximum plasma frequencies of 3 MHz and 10 MHz in the E and F region, respectively, and a radar frequency of 14 MHz. Rays are plotted for elevation between 0 deg. and 30 deg. Horizontal and vertical axes having different scales, angles are not conserved on this plot, while the earth's curvature induces the curvature seen on low elevation rays. Small triangles are plotted along a ray path every 100 km, and thicker lines are plotted when the aspect sensitivity condition is met within 1 deg. On Fig. 4, we notice that scatter appears below 100 km in the E region and between 200 and 300 km in the F region. Echoes returned from these altitudes will be distinguished by their range, 600 to 800 km for F region echoes and 300 to 500 km for E region echoes.

The variations which can be expected when changing the radar frequency are plotted on Fig.5 for a fixed elevation angle. Waves are reflected at higher altitudes when the radar frequency increases. In the present example, waves with frequencies below 15 MHz are reflected in the E region and do not reach perpendicularity to the magnetic field. Higher frequency waves reach the F region where the can detect field aligned scatter. One also notices that the range of the scatter slightly increases with increasing frequency. As a further step, we have included latitudinal gradients of electron density, modelling blobs commonly present at high latitudes (Fig. 6). The structure plotted in Fig. 6a has a density increase of 100% over background. Its width is 150 km and it can be located at various distances from the transmitter. Blobs with lower density (less than 50%) do not induce noticeable changes in the ray paths. The associated ray tracing (Fig. 6b) shows that the region of near-perpendicularity in the F region is extended, both in altitude and range, compared to previous cases. Results will evidently vary according to the density model, but scatter region can, in principle, be localized when one knows the density profile, both in latitude and altitude. This can only be obtained with an incoherent scatter radar operating in a meridional scanning mode. Such data are gathered only during specific periods, while the HF coherent radar performs observations during more extended periods. Two-dimensional mapping of the convective flow will therefore be done by combining radial velocities with the hypothesis of a straight line propagation. Previous studies have concluded that the range drift is of the order of a radar resolution cell (Villain et al., 1984). The simple ray tracing presented here confirms this conclusion, and also show that E and F region echoes will usually be separated in range, and that a larger spatial coverage of the scattering region can be obtained by using several radar frequencies separated by a few MHz.

## EXPERIMENTAL DATA

The first joint observations of the Goose Bay and Schefferville radars were performed in october 1986. The Goose Bay radar was operating in the usual azimuthal scanning mode on 16 beams with a temporal resolution of 80s to complete a map. In Schefferville, the radar was stepping in frequency from 8 Mhz to 18 MHz by 2 MHz, with a temporal resolution of 20s on each frequency. The antenna beam was fixed on a direction 15° east of geographic north.

The temporal variation of backscattered power and radial velocities measured in Schefferville during 4 hours is plotted on Fig. 7 for several radar frequencies. Several types of scatter are present on these figures. At shortest ranges, from 300 km, scatter is detected at the same distance for all

frequencies. At such a short range, and due to its lack of variation with radar frequency, this scatter can be associated with E region irregularities. Radial velocities are globally positive (towards the radar) until 01:30 UT (LT = UT-4) and reverse afterwards. At this local time, the radar has therefore certainly detected the velocity reversal associated with the Harang discontinuity.

A second region of scatter is present from 00:00 UT until 01:00 UT, but only on the lowest frequencies (8 MHz to 12 MHz). The range is increased compared to E region scatter, and varies with radar frequency, from around 900 km at 8 MHz up to more than 1200 km at 12 MHz. It is therefore originating in the ionospheric F region. As shown in the preceeding section, the range variation and the absence of radar return at higher frequencies is easily explained by the effects of refraction. Its effect on the ray is more important at lower frequencies : bending is more pronounced and the normality condition is met at shorter ranges when radar frequency is decreased. At higher frequencies (14 MHz in the present case), the ray is not sufficiently bent to reach perpendicularity with the magnetic field, and the radar cannot detect irregularities, even if they are present in the ionosphere. Observing the F region scatter as seen on Fig. 7a to 7c, one immediately concludes that the frequency variation will be usefull to extend the spatial coverage of the system. On 8 MHz, echoes are seen from 700 km to 1200 km, while the shortest and furthest ranges are 900 km or 1100 km and 1500 km or 1600 km on 10 MHz and 12 MHz, respectively. Merging the three frequencies, the radar covers the region between 700 km and 1800 km, or approximately 10 degrees in latitude. When combining the data from the two radars, one will therefore derive extended maps of the plasma convection.

An example of power and velocity maps observed at 00:00:40 UT with the Goose Bay radar are shown on Fig. 8a and 8b, respectively. They show the presence of a scatter region over an extended azimuthal sector (10 of the 16 antenna beams, or 35 deg.) with a maximum S/N ratio of 20 dB. The radial velocity is directed towards the radar, with values of the order of 600 m/s in the westward part of the map, and values near zero or slightly away from the radar when looking towards north. The radial velocities measured with the Goose Bay radar have been combined with those observed in Schefferville to derive two-dimensional velocities in common points, and have been extended to the whole Goose Bay map with the flux divergence-free method described earlied. A sample map is shown on Fig. 9, for two different frequencies of the Schefferville radar, namely 8.8 MHz (Fig. 9a) and 10.7 MHz (Fig. 9b).

The resulting convection pattern is similar on the two maps. This is not surprising, as the data used for each derivation are not totally independent : the Goose Bay map of radial velocities is the same, and only the Schefferville values have been changed between them. The global pattern shows an equatorward motion of the plasma, with an eastward rotation of the flow at lower latitudes. Some westward components are also visible on the westward edge of the map. This convection pattern is in agreement with expectations at this particular local time : the plasma flows out the polar cap into the late evening auroral zone, and turns to a westward flow in the pre-midnight period, turning eastward in the post-midnight period with a transition at the Harang discontinuity. The two maps cover about 5 degrees in latitude at both frequencies, but one can note that the second one is slightly displaced northwards compared to the first one. Multifrequency operations of the radar is therefore, as mentioned previously, an important point in obtaining a spatial coverage as extensive as possible.

The high temporal resolution of the radar makes it possible to map the convection pattern every 80s. Fig. 10 shows 3 successive maps computed at this interval, each of them obtained for a different frequency of the Schefferville radar. One notes the variation with frequency in spatial coverage and geographical location, both being a consequence of the change in the ionospheric propagation. It is therefore important to adjust the radar frequency as to optimize the coverage of the map. The sequence shows, as does Fig. 9, a convective flow out of the polar cap which rotates eastward at lower latitudes.

**CONCLUSION**

Mapping the plasma convection at the ionospheric footprints of the most external regions of the magnetosphere is one of the present research goals in ionospheric physics. The data obtained with the two radars of the PRISM program show that HF coherents radars can play an useful role in this research effort, in coordination with other ground-based and space-borne instruments. Each radar measures the Doppler velocity of F region ionospheric irregularities, which is directly related to the ion drift. Combining radial velocities observed by two radars leads then directly to two-dimensional mapping of the plasma flow. Another advantage of the radars is the large spatial coverage, compared to VHF radars operating in the E region. For example, the maps shown in Fig. 9 extend over 5 deg. in latitude, and this coverage represents only a small part of the potential radar field of view. By comparison, the maximum coverage of the STARE radar is only of the order of 4 deg. in latitude (Greenwald et al., 1978). Temporal and spatial resolutions are of the same order of magnitude for HF or VHF/UHF radars. Refraction of the HF radio wave in the ionosphere is both a benefit and an handicap. On one side, it allows the radar to detect scatter in the very high latitude regions, but, on the other side, it induces some uncertainty in the geographical location of the target. Effects are nevertheless minimal when dealing with studies of large scale patterns.

The results presented in this paper were obtained with the Goose Bay radar scanning over 50° in azimuth and the Schefferville radar transmitting in a fixed direction. We have therefore been obliged to use the hypothesis of a divergence-free flow to map the plasma convection. An immediate improvement to the system is an extension of the scanning capability of the Schefferville radar. This will be set up during fall 1988. All velocity values will thereafter be derived from simultaneously measured radial components at the same point. Studies of plasma convection will also benefit from other improvements already installed. Firstly, the Goose Bay radar possess the capability of measuring vertical angle of arrival of the backscattered signal, owing to an interferometry mode using a second parallel antenna array in front of the main array. Secondly, another HF coherent radar has been set up in Halley Bay (Antartica) by the Bristish Antartic Survey and APL. The program, known as PACE (Polar Anglo-American Conjugate Experiment) is aimed to study the ionospheric convection at conjugate points. With this set of three coherent HF radars, one can therefore hope to gather useful data for theoretical and numerical works on the ionospheric and magnetospheric physics at very high latitudes.

**REFERENCES**

BAKER K.B., GREENWALD R.A., WALKER A.D.M., BYTHROW P.F., ZANETTI L.J., POTEMRA T.A., HARDY D.A., RICH F.J., RINO C.L., A case study of plasma processes in the dayside cleft, J. Geophys. Res., 91, 3130-3144, 1986.

BANKS P.M., DOUPNIK J.R., A review of auroral zone electrodynamics deduced from incoherent scatter radar observations, J. Atmos. Terr. Phys., 37, 951-972, 1975.

CHATURVEDI P.K., OSSAKOW S.L., Nonlinear stabilization of the current convective instability in the diffuse aurora, Geophys. Res. Lett., 6, 957-959, 1979.

CROCHET M., BARPEAU D., De MAISTRE J.C., Analyse de la propagation par rétrodiffusion directe sur les irrégularités alignées dans la couche F aurorale. Influence d'une perturbation, Ann. Geophys., 27, 453-461, 1971.

GREENWALD R.A., WEISS W., NIELSEN E., THOMSON N.R., STARE : A new auroral backscatter experiment in northern Scandinavia, Radio Sci., 13, 1021-1039, 1978.

GREENWALD R.A., BAKER K.B., VILLAIN J.P., Initial studies of small-scale F-region irregularities at very high latitudes, Radio Sci., 18, 1122-1132, 1983.

GREENWALD R.A., BAKER K.B., HUTCHINS R.A., HANUISE C., An HF phased-array radar for studying small-scale structure in the high-latitude ionosphere, Radio Sci., 20, 63-79, 1985.

HANUISE C., VILLAIN J.P., CROCHET M., Spectral studies of F-region irregularities in the auroral zone, Geophys. Res. Lett., 8, 1083-1086, 1981.

HANUISE C., GREENWALD R.A., BAKER K.B., Drift motions of very high latitude F region irregularities : Azimuthal Doppler analysis, J. Geophys. Res., 90, 9717-9725, 1985.

HOLT J.M., WAND R.H., EVANS J.V., Millstone Hill measurements on 26 February 1979 during the solar eclipse and formation of a midday F-region trough, J. Atmos. Terr. Phys., 46, 251-258, 1984.

HUNSUCKER R.D., Analysis of a backscatter signature obtained with a high-resolution HF radar system, Radio Sci., 6, 763-768, 1971.

JONES R.M., STEPHENSON J.J., A versatile three-dimensional ray tracing program for radio waves in the ionosphere, Rep. 75-76, Office of Telecom., US Dept. of Commerce, Boulder, Co., 1975.

KESKINEN M.J., OSSAKOW S.L., Mc DONALD B.E., Nonlinear evolution of diffuse auroral F region ionospheric irregularities, Geophys. Res. Lett., 7, 573-576, 1980.

MILLMAN H.G., Field-aligned ionization scatter geometry, J. Geophys. Res., 74, 900-905, 1969.

MILLMAN H.G., Refraction effects on magnetic field geometry and HF propagation, J. Atmos. Terr. Phys., 37, 751-760, 1975.

NIELSEN E., SCHLEGEL K., Coherent radar Doppler measurements and their relationship to the ionospheric electron drift velocity, J. Geophys. Res., 90, 3498-3504, 1985.

RUOHONIEMI J.M., GREENWALD R.A., BAKER K.B., VILLAIN J.P., Mc CREADY M.A., Drift motions of small-scale irregularities in the high latitude F region : an experimental comparison with plasma drift motions, J. Geophys. Res., 92, 4553-4564, 1987.

SWENSON E.M., Aspect. sensitive reflections from ionization irregularities in the F-region, J. Atmos. Terr. Phys., 34, 1469-1476, 1972.

VILLAIN J.P., GREENWALD R.A., VICKREY J.F., HF ray tracing at high latitudes using measured meridional electron density distributions, Radio Sci., 19, 359-374, 1984.

VILLAIN J.P., CAUDAL G., HANUISE C., A SAFARI-EISCAT comparizon between the velocity of F-region small-scale irregularities and the ion drift, J. Geophys. Res., 90, 8433-8443, 1985.

VILLAIN J.P., GREENWALD R.A., BAKER K.B., RUOHONIEMI J.M., HF radar observations of E region plasma irregularities produced by oblique electron streaming, J. Geophys. Res., 92, 12327-12342, 1987.

**ACKNOWLEDGEMENTS**

The SHERPA radar in Schefferville is supported by Institut National des Sciences de l'Univers (INSU). This study has also been partly funded by Direction des Recherches et Etudes Techniques (DRET). We wish to thank the personnel of Mc Gill Subartic Research Station for daily operation of the radar at Schefferville. The Applied Physics Laboratory HF radar at Goose Bay 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 grant ATM-8506851, in part by the National Aeronautics and Space Administration (NASA) under NASA grant 7055, and in part by the Defense Nuclear Agency and Rome Air Development Center under contract N00039-87-C-5301. We would like to thank J. Kelsey and his co-workers for help in day-to-day operation of the radar and the Ionospheric Effects Branch of the Air Force Geophysics Laboratory for permission and support in operating from its Goose Bay field site.

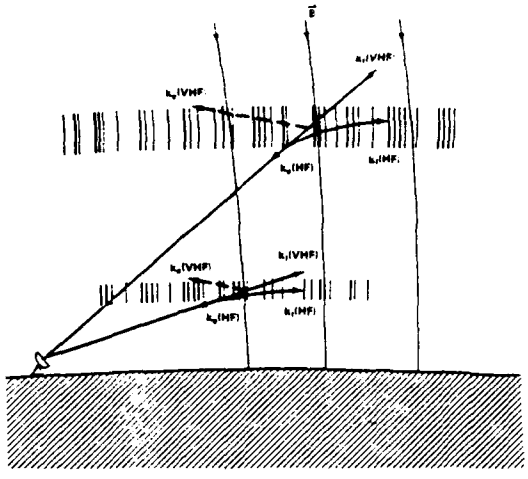


Fig. 1. Diagram illustrating manner in which VHF and HF radar signals are scattered into space by very high latitude E and F region irregularities. At high frequencies, the radar signals are refracted toward the horizontal and the scattered signals may return to the radar.




Fig. 2. Field of view of the two PRISM HF radars. The JHU/APL radar, located at Goose Bay, Labrador, can be steered electronically into one of possible 16 directions. The SHERPA radar in Schefferville is presently directed towards one direction, but a similar scanning capability will be added soon.

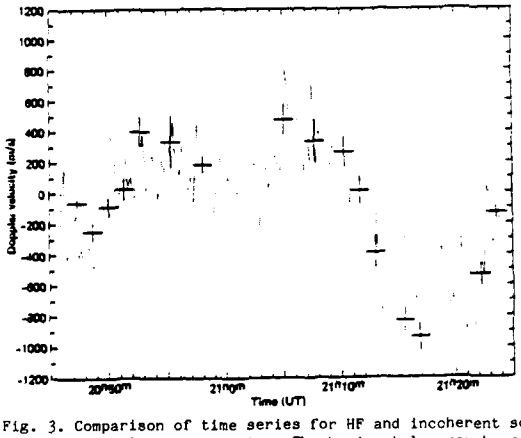



Fig. 3. Comparison of time series for HF and incoherent scatter velocity estimates. The horizontal error bars on the HF data are shown exaggerated by a factor of 10. (from Ruohoniemi et al., 1987).

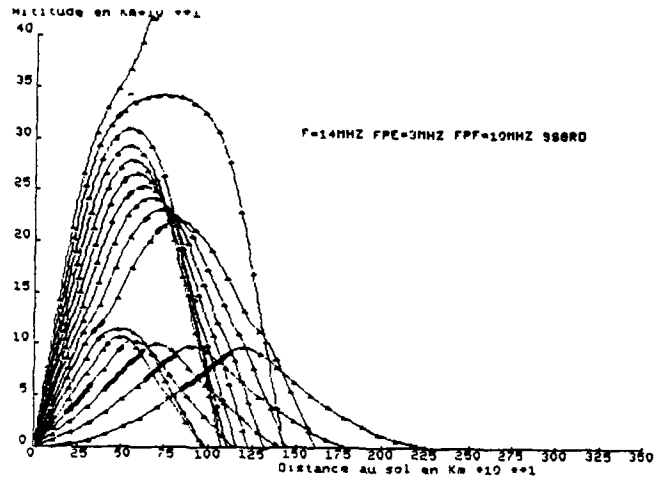



Fig. 4. Ray tracing showing the propagation of an HF wave according to the elevation of transmission. The ionosphere is modelled with 2 parabolic layers of maximum plasma frequency 3 MHz and 10 MHz, respectively. The radar frequency is 14 MHz. Thicker lines indicate the points where the ray is nearly perpendicular to the earth's magnetic field. Scatter can be detected in both E and F regions.

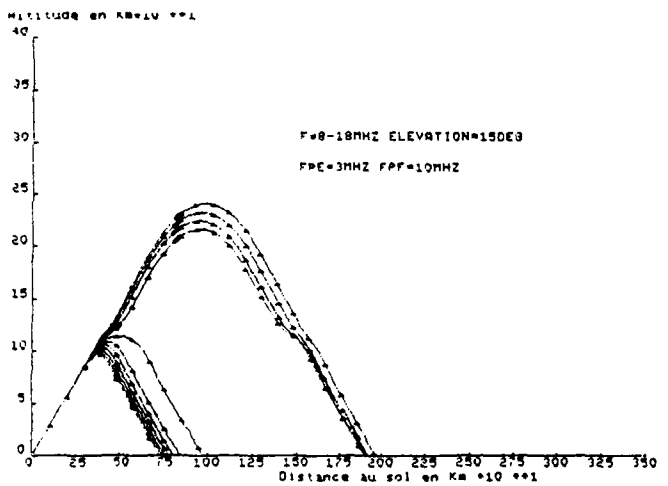
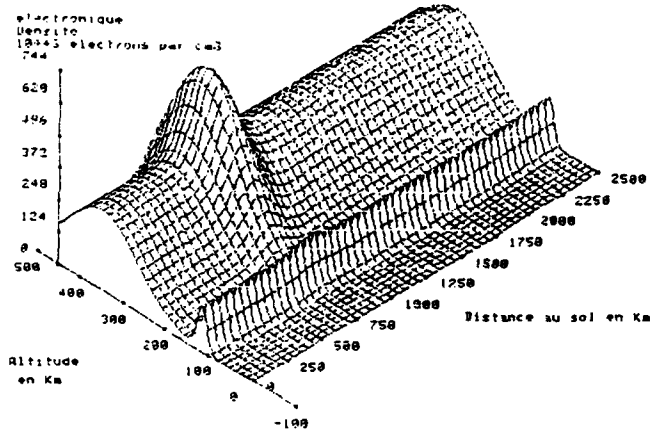



Fig. 5. Ray tracing with the same parameters as in fig. 4, but showing the variation with frequency for a fixed elevation angle. In this example, lower frequencies are reflected in the E region. Scatter can be detected in the F region at higher frequencies, at a range increasing with frequency.



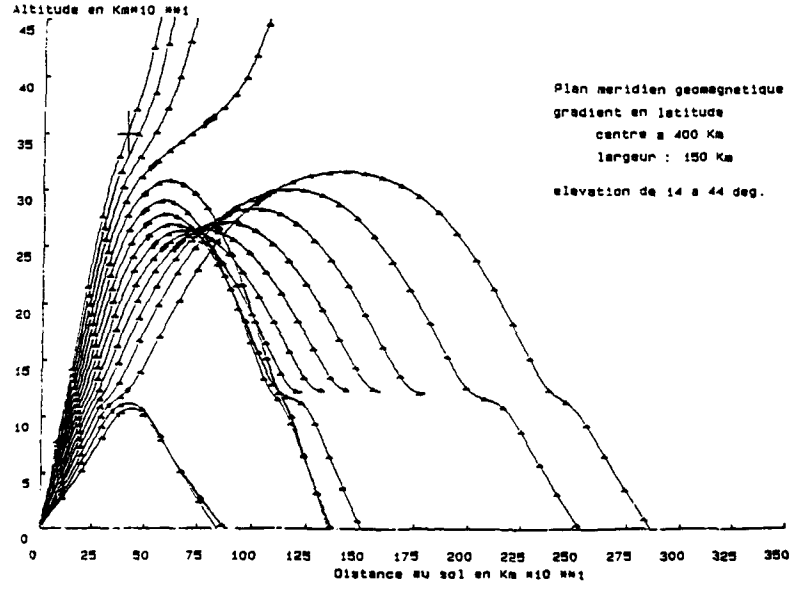




Fig. 6. Ray tracing in an irregular ionosphere where latitudinal density gradients are present. The ionospheric model includes a density blob with an increase of 100 % over background (fig. 6a). The resulting propagation paths show that F-region irregularities can be detected between 250 km and 350 km altitude.

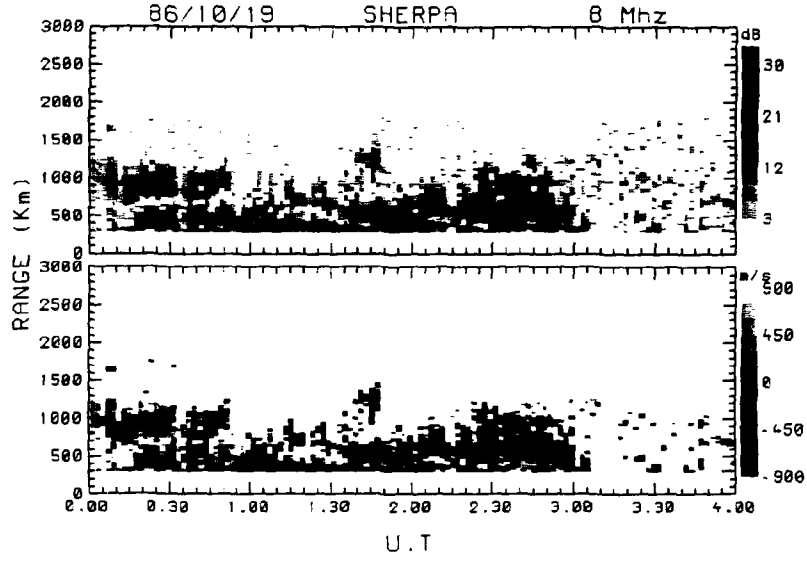



Fig. 7a.

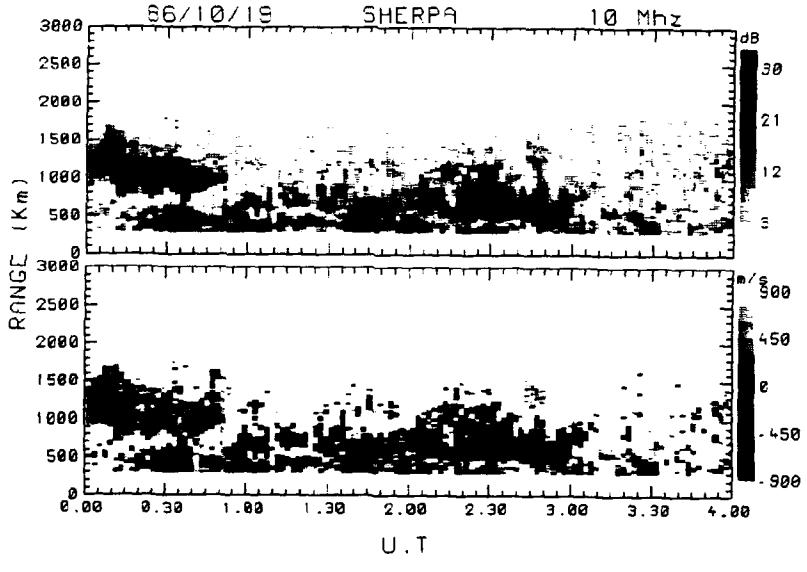



Fig. 7. Temporal variation of the backscattered power and Doppler velocity observed at Scheffferville on 19/10/86 at frequencies from 8 MHz to 14 MHz. E region scatter is present at shorter ranges. F region scatter is also detected from 00:00 UT to 01:00 UT up to 12 MHz. The range of the F region echoes increases with increasing frequency, therefore extending the radar spatial coverage.

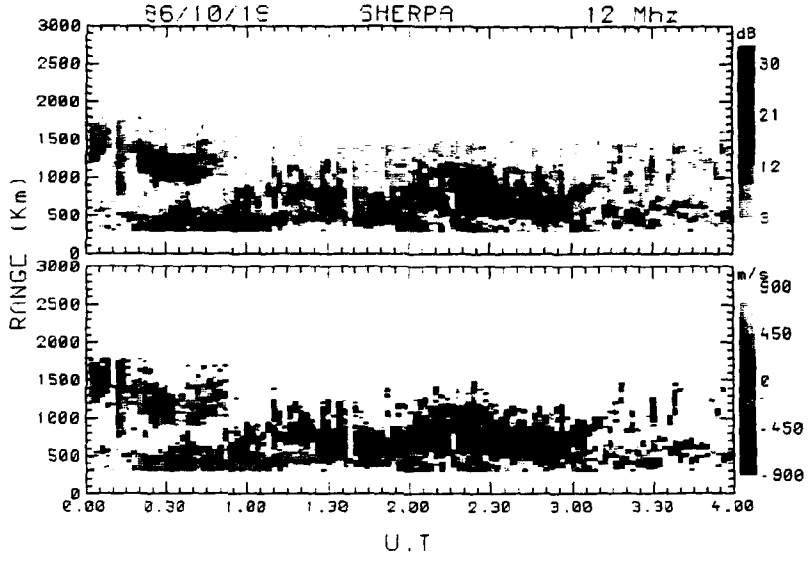



Fig. 7c.

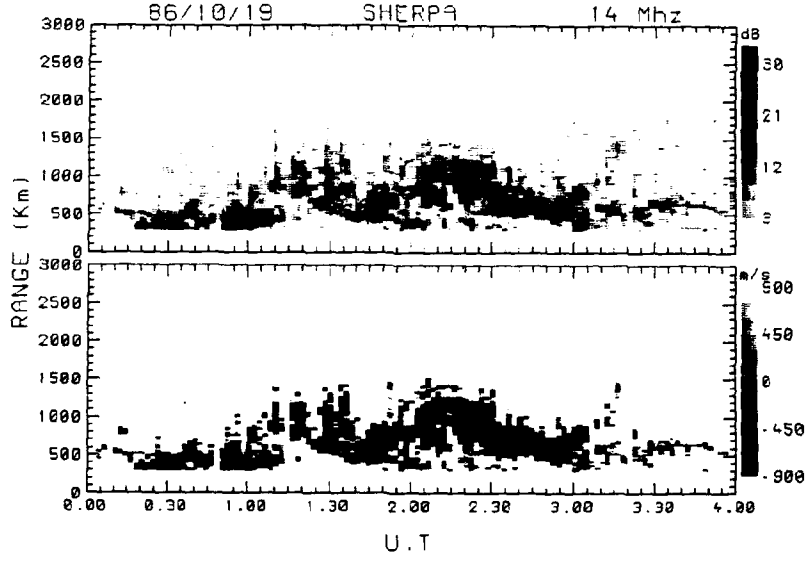



Fig. 7d. F region echoes are not detected at 14 MHz, as the ray is not sufficiently refracted towards the horizontal to reach perpendicularity to the magnetic field.

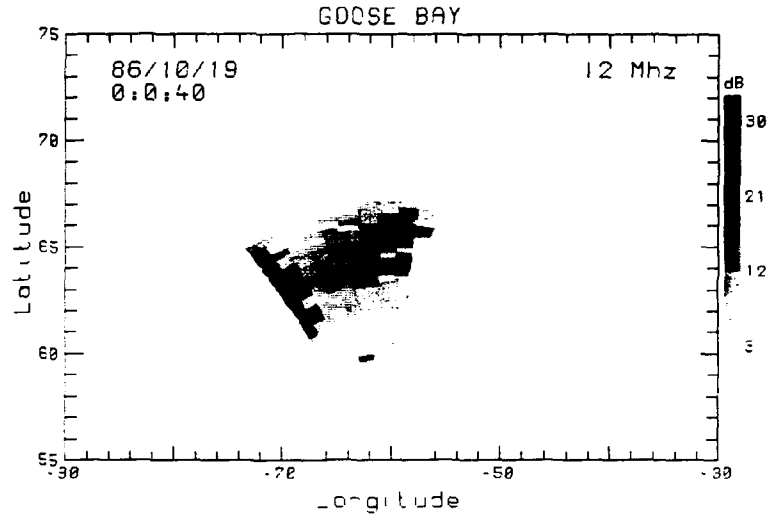



Fig. 8a.

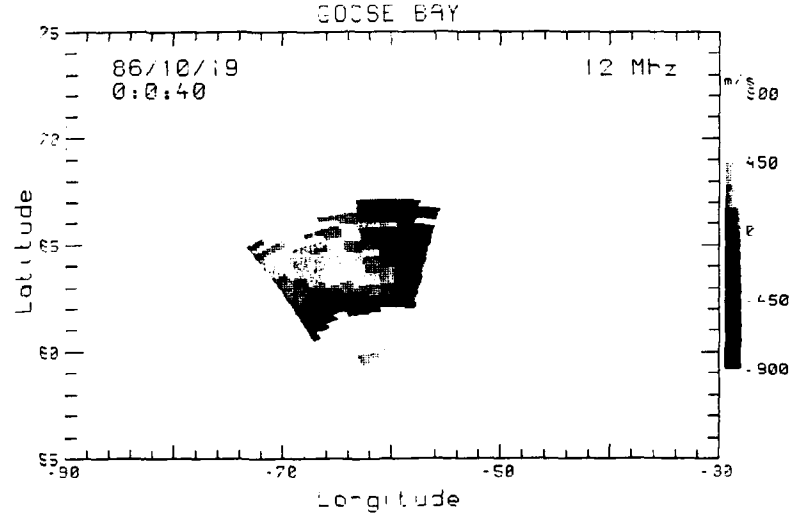



Fig. 8. Map of backscattered power (fig. 8a) and radial velocity (fig. 8b) for the Goose Bay radar on 19/10/86. Maximum signal to noise ratio reaches 20 dB. Velocities are directed towards the radar with values up to 600 m/s.

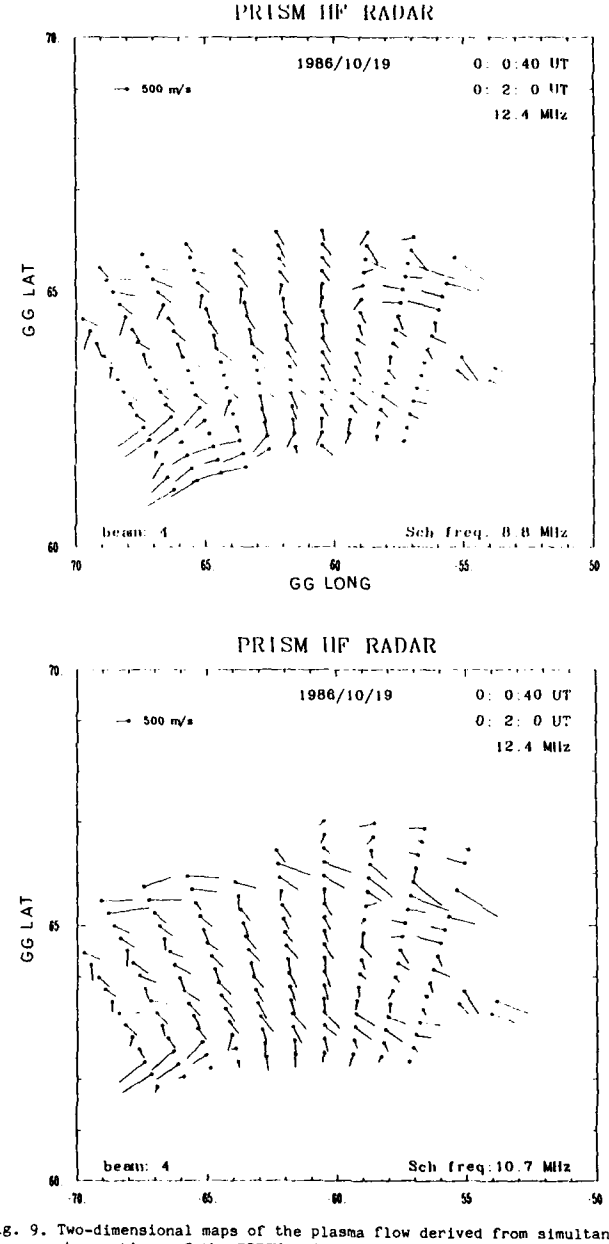



Fig. 9. Two-dimensional maps of the plasma flow derived from simultaneous observations of the PRISM radars. The two maps are obtained for the same Goose Bay data but different SHERPA frequencies, 8.8 MHz (Fig. 9a) and 10.7 MHz (Fig. 9b). The patterns are consistent with a flow coming out of the polar cap and rotating eastward at lower latitudes.

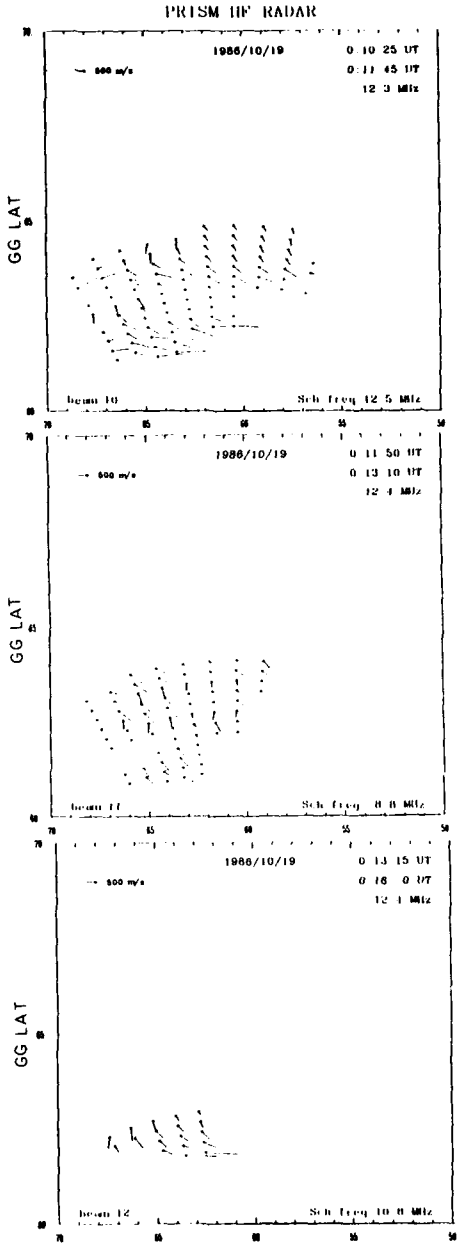



Fig. 10. Example of temporal variation of the flow pattern with a resolution of 80 s. The pattern is consistent with fig. 9. The change in frequency between the 3 snapshots shows the importance of optimizing the radar operation and the possibility of extending the spatial coverage by mixing several frequencies.

## DISCUSSION

**C.Haldoupis**

You mentioned that you receive echoes with the HF radars from electrojet $E$ region (100–120 km) and $F$-region (200–300 km). Do you have any indication of receiving echoes from the upper $E$ region in the altitude range around 130–160 km, and if yes, under what geophysical conditions do they occur and what are their characteristics? I am asking this having in mind that the upper region is susceptible to the ion cyclotron instability with EIC wavelengths in the range you use (i.e., 10–20 km), and therefore, should be detectable by your radar system.

**Author's Reply**

Echoes detected with HF radars come from the $F$ region, generated by gradient-drift mechanisms, from electrojets, generated by both two-stream and gradient drift instabilities, and also from the upper $E$-region. In this later case, observations performed with the PRISM radars are consistent with, on one side ion-cyclotron waves and, on the other side, ion-acoustic waves with a destabilizing effect of field-aligned currents. Studies are presently in progress on these two processes.

**J.Belrose**

In your presentation you spoke about the possibility to observe HF backscatter from irregularities for a range of distances from short to long. This, of course, depends on the frequency used and the propagation conditions, but also on the antenna system employed which is given little consideration by the HF radar researcher. For the Goose Bay, Labrador HF radar, a log periodic dipole array is used, at 15.2m. For this height, at a frequency of 10 MHz, the launch angle will be about 22°, and the radar will be optimum for observing backscatter at a distance range of about 1250 km. The sensitivity of the system will be less for shorter and longer ranges. Could you comment on the antenna system you employ?

**Author's Reply**

The antenna system in Scheffeville is similar to the one installed in Goose Bay. Both are, therefore, optimized to look at relatively low elevation angles (10°–30°), to avoid direct reflections from the ionosphere. Echoes are observed typically at ranges between 400 and 800 km for the $E$ region (around 100 km) and at ranges from 1000 km to 2000 km for the $F$ region echoes. This depends on the radar frequency, or the presence of irregularities and on the ionospheric plasma frequency.

**K.Bibl**

I don't know how good your ground screen is in Goose Bay, but Goose Bay has the lowest conductivity in the whole world.

**Author's Reply**

Yes. I think there was a ground screen put in Goose Bay.

# EXAMPLES OF METEOROLOGICAL BEHAVIOR OF THE IONOSPHERE

K. Bibl, University of Lowell, Center for Atmospheric Research,
450 Aiken Street, Lowell, MA 01854, USA

## SUMMARY

For some time the network of ionosonde stations in Europe had been dense enough to study the meteorological behavior of the ionosphere. Because the electron recombination near the maximum of the F-region ionization is sufficiently small, the maximum electron density and the profile are controlled substantially by dynamic processes. Gravity waves and vertical plasma drift and its vertical gradients change the F-layer ionization significantly. In fixed frequency ionosonde recordings the author has discovered direct oblique echoes from a hole in the ionosphere over the Alps. Maps have been drawn to show the meteorological properties of this "European Anomaly" either as a time development of a cross-section of Europe from northern Germany to Rome, Italy using six stations, or as a sequence of contour maps of the ionosphere over Europe using ten stations. A short animated movie was produced demonstrating the development of the hole in the ionosphere and its recovery with a curl developing.

Even the average behavior of the F-region ionization shows substantial differences with location. In Europe the variations of the local gradients in ionization can be different by a factor of two over two locations separated by 1000 km.

This behavior, important for understanding the meteorology of the ionosphere and for precise ionospheric radio predictions, certainly requires intense studies with digital sounders in Europe and at other locations.

## INTRODUCTION

In the study of the dynamic behavior of the ionospheric F region three types of important events must be separated:

(1) acoustic-gravity waves travelling equatorwards rather horizontally, parallel to the magnetic declination

(2) magnetic substorm effects blowing the F-region ionization up and away in the magnetic field tubes and

(3) meteorological phenomena which produce local extrema in the ionospheric F-region ionization. Only the meteorological type of local ionospheric disturbances shall be described in this report.

There are two reasons why these events have not been studied more intensely in the last 25 years.:

a) the network of ionosonde stations is not dense enough and

b) the rules of ionogram analysis tend to fill in temporary minima in the F region parameters. Extrapolation of foF2 to infinite delays in case of "absorption" near the critical frequency of the F-region often extends an oblique incidence trace, hiding a much lower foF2 overhead. Figure 1 shows a sequence of Digital ionograms during a local temporary minimum in F2-layer ionization. If it were a

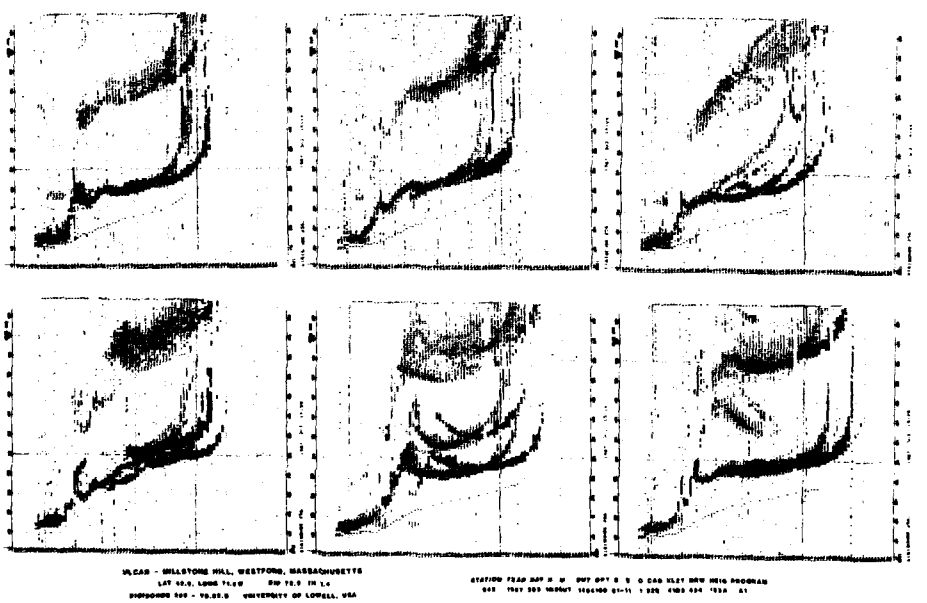


Figure 1 Quarter-Hour Sequence of Digital Ionograms showing echo amplitude, polarization, and obliquity.

*Traveling Ionospheric Disturbance* (TID) caused by an acoustic gravity wave, it would be out of the antenna beam within 20 minutes. On the contrary, such an event could easily be the source of a gravity wave. Records of the new digital ionosondes confirm those findings of 25 years ago. The beam of the receiving antenna array is switched from pulse to pulse into six different azimuths at a specific zenith angle in addition to the overhead position where echo polarization is identified. By these means more ionospheric holes are discovered and their location and motion followed.

## THE EUROPEAN ANOMALY

For many years the ionospheric station in Breisach, Germany, had produced continuous multi-directional color records of ground backscatter echoes via the ionosphere on several fixed frequencies. In these records the author found direct echoes from the F-region at a frequency much higher than the vertical critical frequency. Because the echoes come only from one specific direction it was assumed first that the ionization was much higher in the south than overhead. But an intensified study of the F-layer critical frequency during the times of those events showed the contrary: an ionization minimum south of the Breisach station. Figure 2 shows 44 maps with time as abscissa and six stations at their respective location on the ordinate: L for Lindau (near Goettingen), Germany; D for Dourbes in the south of Belgium; F for Breisach (near Freiburg), Germany; S for Sottens in Switzerland; G for Genoa, Italy; and R for Rome, Italy.

The events are substantial: the ratio of electron densities between maxima and minima reaches sometimes a value of 2 or more (Bibl, 1964A). With these maps the ionospheric station in Switzerland was vindicated; it was suspected before that this station produced erroneous data. But it is obvious that these deep

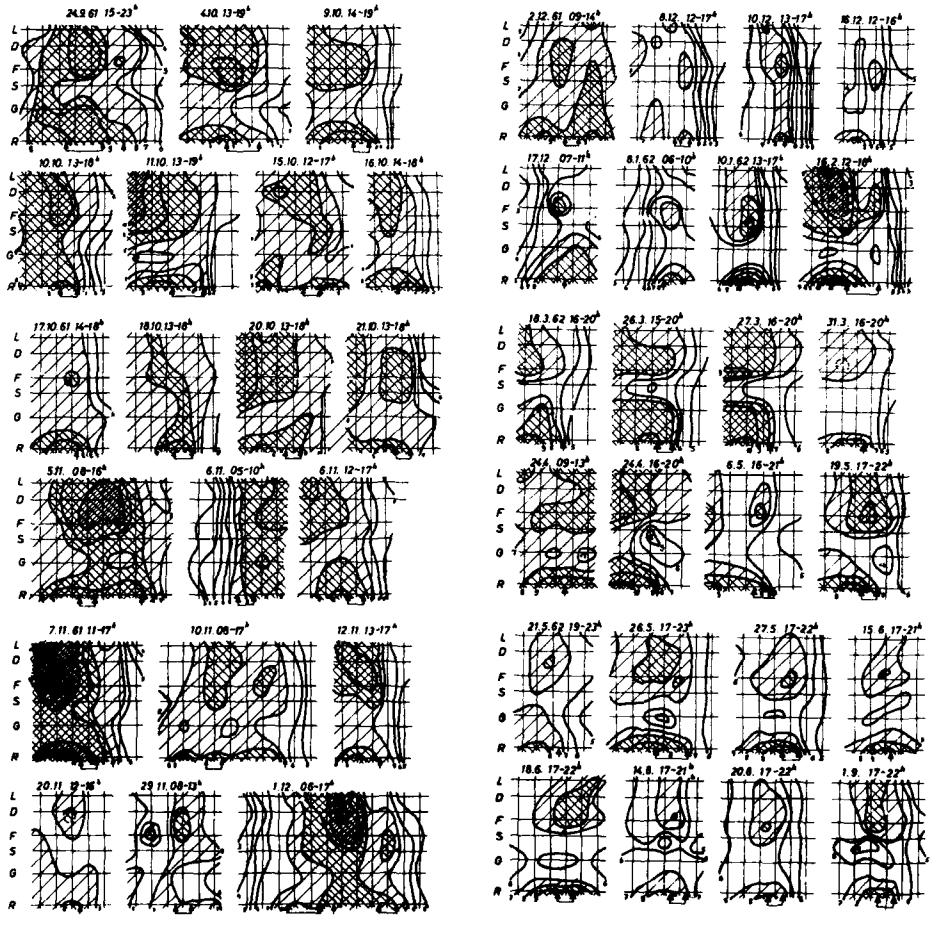


Figure 2 Time Sequence of a Cross-Section Through Europe During Oblique Reflection Events

minima will even effect the median values, as Figure 3 shows. Only about half of the cases found are presented here because of missing data at one or several stations being used. It is easily possible that the events are more frequent because they have been selected solely when backscatter recordings showed direct ionospheric echoes above the vertical critical frequency. Night-time events were never found since the sounding frequencies were too high to produce direct oblique echoes at night. They are, however, less likely because they do not show up in the average maps of Figure 3. The day-time minima over Genoa are very clear in October and November, while the minimum is over Sottens in September and December.

LOCAL GRADIENTS AND VARIATIONS OF foF2

For each of six months in 1961 the average of all hours of the day has been put on a two dimensional map of Europe, as Figure 4 shows. The following stations were used: I=Inverness, Scotland; E Slough, England; J Juliusruh, L=Lindau, F=Freiburg, Germany; D=Dourbes, Belgium; S=Sottens, Switzerland; C=Prague, Czechoslovakia; O Graz, Austria; G=Genoa, R=Rome, Italy. Lines of constant foF2 are in 0.1 MHz. In March the gradients between Sottens and Genoa are largest; in October the gradients between Genoa and Rome are largest; in December the gradients between Dourbes and Freiburg are largest.

Even more surprising is the variation of the gradients over Europe. To create a significant parameter the hourly value of each day has been subtracted from the monthly average; then the difference of those DfoF2 values between two adjacent stations has been formed. In the monthly distribution of these differences $Dx,y = DfoF2_x - DfoF2_y$ for station pair x and y, the upper and lower quartiles have been found and divided by the distance between the stations x and y. This quantity shall be called the variation of the gradient. Figure 5 shows a map with lines of equal value for this parameter. Although the higher station density in the south must enhance the local gradient variations because it makes smaller scale variations visible, it is obvious that the ionosphere over northern Europe is substantially quieter in space and time than the ionosphere over the Alps.

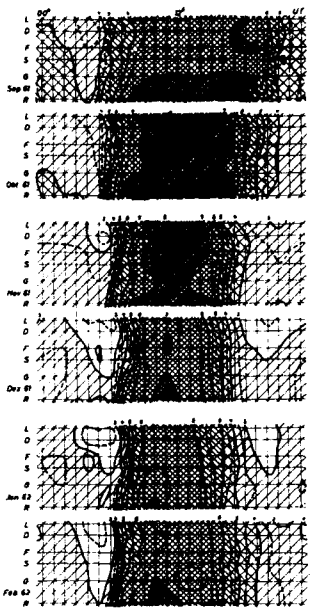


Figure 3 Time Dependence of 6 Monthly Averages for the foF2 cross-section over Europe

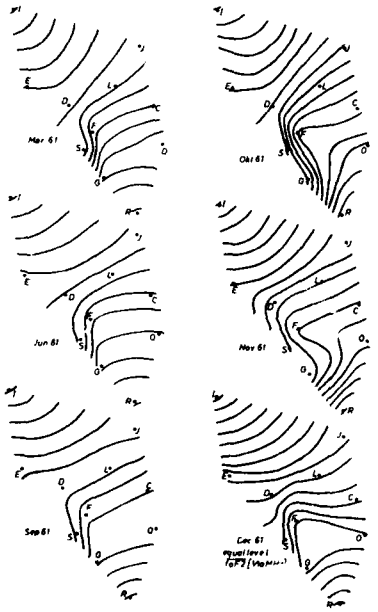


Figure 4 Isodensity lines of foF2 over Europe

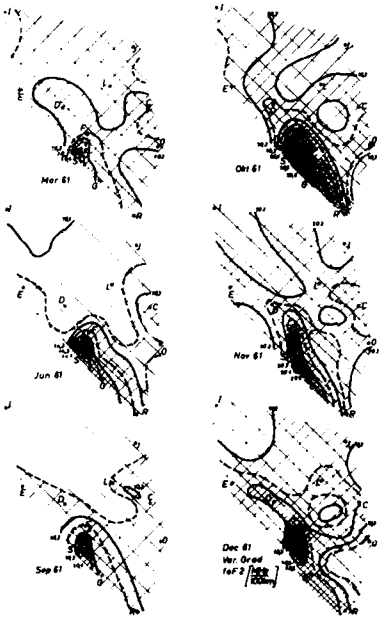


Figure 5 Variation of the Gradients in foF2 over Europe

## MAPS OF METEOROLOGICAL BEHAVIOR

*Using ten* ionospheric stations, 5- and 6-hour sequences of synoptic ionospheric maps have been drawn for four oblique ionospheric echo events. (The events were chosen for cases *when* all stations had complete data.) Ten ionospheric stations were used: L=Lindau, Germany; D=Dourbes, Belgium; C=Prague, Czechoslovakia: F=Freiburg (Breisach), Germany; O=Graz (Oesterreich), Austria: P=Poitier, France; S=Sottens, Switzerland; B=Belgrade, Yuogslavia; G=Genoa, Italy; R=Rome, Italy. In all cases there was a simultaneous enhancement of the local maxima as well as the local minima in the middle of the event. This led to the substantial increases *in* gradients, indicated by the numbers in the top left corner of the individual maps. A second common feature seems to be a counter-clockwise curl of ionization around the minima, well visible in columns 2 and 3 of Figure 6. An animated movie has been made of one of the cases to better show the time development.

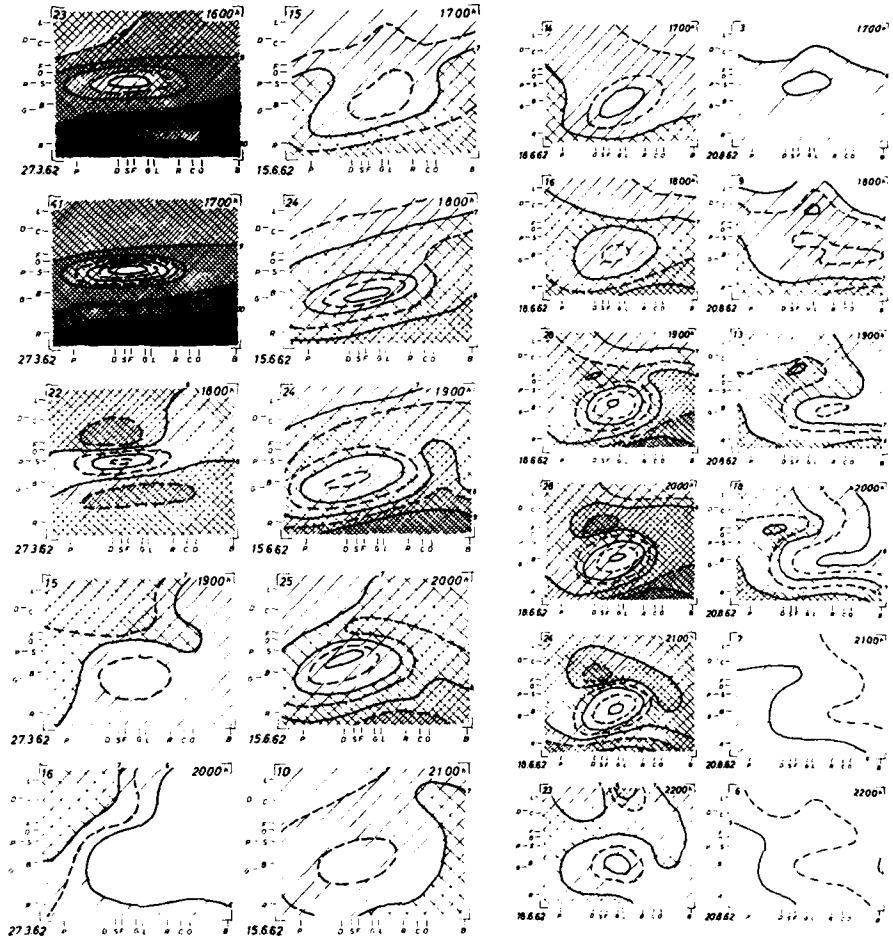


Figure 6 Development of foF2 Over Europe During Four Meteorological Events

## EVENTS AT OTHER LOCATIONS

With the exception of Russian satellite electron density probe recordings which show localized holes in the ionosphere, most researchers have been following TID events. The localized events tend to mess up the picture of the TID studies substantially. In comparing continuous vertical incidence recordings at stations in Belgium, South Germany and Greece, practically no TID event can be followed over 2000 km distance (Bibl, 1962), while motions along the equator show good correlation over larger distances (Bibl 1964b).

Sudden minima in F-layer ionization have certainly been reported at many stations (for example: X.M. Khan et al., 1985), but often these events have been explained as moving from one place to another (Huang and Jeng, 1978). The occurrence of an event at different times for different stations can only be called a motion if data from three stations in a row are available.

From more recent measurements with digital ionosondes it can also be concluded that over North America the local gradients are quite different. For example, the ionosphere over Boulder, Colorado, is much quieter in time and space than the ionosphere over Lowell, Massachusetts. Certainly the distance between the Aurora Zone and the Magnetic Equator is much shorter at the East Coast of the North American continent then on the West Coast, but this does not seem to be the only reason.

It certainly was a shame that two West Coast stations were closed because their data did not fit in the prediction map. Also several European stations were given up because of the high cost of data analysis.

## LIMITED EXPLANATION

It seems impossible that a global wind system in ionospheric height, as described by Rishbeth (1967), could explain such a localized event, but local wind patterns and electric fields extending from the troposphere to the ionosphere could change the vertical plasma drift, modifying the electron density profile and the maximum electron density substantially without changing the total electron density. This could produce enhancements and "bite-outs" simultaneously at two close locations. When the local electric field differences are equalized, an average electron density distribution could then be restored.

## FUTURE POSSIBILITIES

The new chain of digital ionosondes: Qaanaaq, Greenland; Goose Bay, Labrador; Argentia, Newfoundland; Lowell, Massachusetts; and Wallops Island, Virginia; enables us to separate the three phenomena: TIDs, substorms, and meteorological events. We also hope that the European network will be revived to study the meteorology of the ionosphere. Automatic data analysis will make the routine efficient and free manpower for scientific studies. Other countries, like Pakistan and Japan, will also make a substantial effort in this direction.

## ACKNOWLEDGEMENT

Some of the figures were reprinted from Annales de Geophysique.

The original research was sponsored by the Cambridge Research Laboratories Contract No. 61(052)-623. Additional investigations were partially supported under subcontract NWRA-86-S 002 for AFGL.

## REFERENCES

K. Bibl (1964A), "The Detection of an Important Anomaly in the F-Layer Ionization Above Europe," Ann. Geophys. 20. pp. 447-453.

K. Bibl (1962), "Ionospheric Movements Observed Using the Technique of Direct Characteristic Recording," Proc. Int. Conf. on the Ionosphere, London, pp. 358-362.

K. Bibl (1964B), "Equatorial F-layer Ionization Movements," Ninth AGARD Ionospheric Research Committee Symposium, Copenhagen.

Y.N. Huang and B.S. Jeng (1978), "On the Afternoon Bite-Out in the Critical Frequency of the F2-Layer," JATP 40, p. 581.

Z.M. Khan et al. (1985), "On the Cause of Fore-Noon and Post-Noon Bite-Outs in foF2," JATP 47, pp. 719-724.

H. Rishbeth (1967), "The Effect of Winds on the Ionospheric F2-Peak," JATP 29, pp. 225-238.

## DISCUSSION

**H.Poeverlein**

Gassmann once believed that the albedo of cloud covers may have some influence, perhaps due to variable UV reradiation. What do you think of such an influence?

**Author's Reply**

I had studied some relevant weathermaps, but I did not find an obvious correlation. Electric fields or changes of the jet-stream may couple into the $F$-region and induce vertical plasma motions.

**K.Rawer**

During the years 1946 through 1956 (then in the French military prediction service SPIM), we established ionization maps by hand with monthly medians. Since the time we had enough stations in Africa, we found a comparable problem in southern Africa. There was a minimum rather regularly appearing during daylight hours. When you are drawing maps by hand you have always some smoothing tendency. I guess now that a comparable regional anomaly might exist there.

**Author's Reply**

It is certainly important for mapping and prediction to avoid elimination of any stations measurements unless *proven* wrong because local gradients may be as important for some applications as the averages and the time variations.

**P.Vila**

This minimum so deep was it associated with sunset?

**Author's Reply**

No, it was afternoon.

# MULTISTATION/MULTIPARAMETER OBSERVATIONS WITH A NETWORK OF DIGITAL IONOSONDES

B. W. Reinisch,(1) J. Buchau,(2) K. Bibl(1) and G. S. Sales(1)

(1) University of Lowell Center for Atmospheric Research, Lowell, MA 01854, U.S.A.
(2) Air Force Geophysics Laboratory, Hanscom AFB, MA 01731, U.S.A.

Summary

The global network of modern ionosondes generates a data set of ionospheric characteristics that can serve as test bed for the developing ionospheric models. Remote access to each station makes it possible to use real time data for project planning and radio communication tasks. New ionospheric parameters like ionospheric roughness, tilt angles and drift are now available for each Digisonde location.

## 1. Introduction

A new generation of modern ionosondes is now being deployed world-wide. By the end of 1988, a network of some forty of these ionosondes (Reinisch, 1986) will provide a consistent data set of ionospheric parameters that are automatically scaled in real time. The automated stations output the standard ionospheric parameters, the h'(f) traces with amplitudes and Doppler frequencies, and the electron density profiles. All these data can be remotely accessed by telephone links, and they are, in general, archived on half-inch magnetic tape.

New measuring techniques make it possible to determine ionospheric structure and dynamics in a more or less routine way. After completion of an ionogram, the sounder measures ionospheric drift and structure using high resolution Doppler shift and incidence angle observations. High latitude drift observations monitor the polar cap convection pattern, and first results from Millstone Hill (42.6°N, 71.5°W) show that this pattern can also control the drift at this mid-latitude location. F-layer tilts, measured at mid-latitude, show typical tilt angles of 2 to 4°; values of 10° are observed occasionally. Observations of the smoothness of the mid-latitude F-layer, as defined by the inverse of the size of the reflection area, show a characteristic day/night variation. During the day the echoes arrive within a small angular cone of $\theta \leq 5°$ (smooth ionosphere), while this angle increases to some 30° at night (rough ionosphere).

## 2. Digisonde Network

**DIGISONDE STATION LOCATIONS**
March 1988
Operating and Near Term Planned

| STATION | LATITUDE | LONGITUDE |
|---|---|---|
| Qaanaaq, Greenland | N77.5 | E290.8 |
| College, USA | N64.9 | E212.2 |
| Sitka, USA | N57.0 | E224.8 |
| Slough, England | N54.5 | E359.4 |
| Goose Bay, Canada | N53.3 | E299.5 |
| Attu, USA | N52.8 | E186.9 |
| Croughton, England | N52.0 | E358.8 |
| Dourbes, Belgium | N50.1 | E 4.6 |
| Munchen, FRG | N48.2 | E 11.6 |
| Argentia, Canada | N47.6 | E307.3 |
| Camp Darby, Italy | N43.5 | E 10.3 |
| Lowell, USA | N42.6 | E288.5 |
| Roquetes, Spain | N40.8 | E 0.3 |
| Beijing, China | N39.9 | E116.5 |
| Wallops Is., USA | N37.9 | E284.5 |
| Diyarbakir, Turkey | N37.9 | E 40.2 |
| Sao Miguel Is., Azores | N37.5 | E334.5 |
| Kunsan, S. Korea | N36.0 | E126.6 |
| Kokubbunji, Japan | N35.7 | E139.5 |
| Xinxiang, China | N35.3 | E113.9 |
| Vandenberg, USA | N34.7 | E239.4 |
| Islamabad, Pakistan | N33.8 | E 72.9 |
| Hamilton, Bermuda | N32.4 | E295.3 |
| Central Texas, USA* | N29.4 | E261.7 |
| Patrick, USA | N28.2 | E279.4 |
| Karachi, Pakistan | N24.8 | E 67.1 |
| Maui, USA | N20.5 | E203.7 |
| Bonnquen, Puerto Rico | N18.5 | E292.9 |
| Baguio, Phillipines | N16.3 | E120.6 |
| Learmonth, UK | S22.1 | E114.0 |
| Sao Paulo, Brazil | S23.5 | E313.5 |
| La Trobe, Australia | S37.8 | E145.0 |

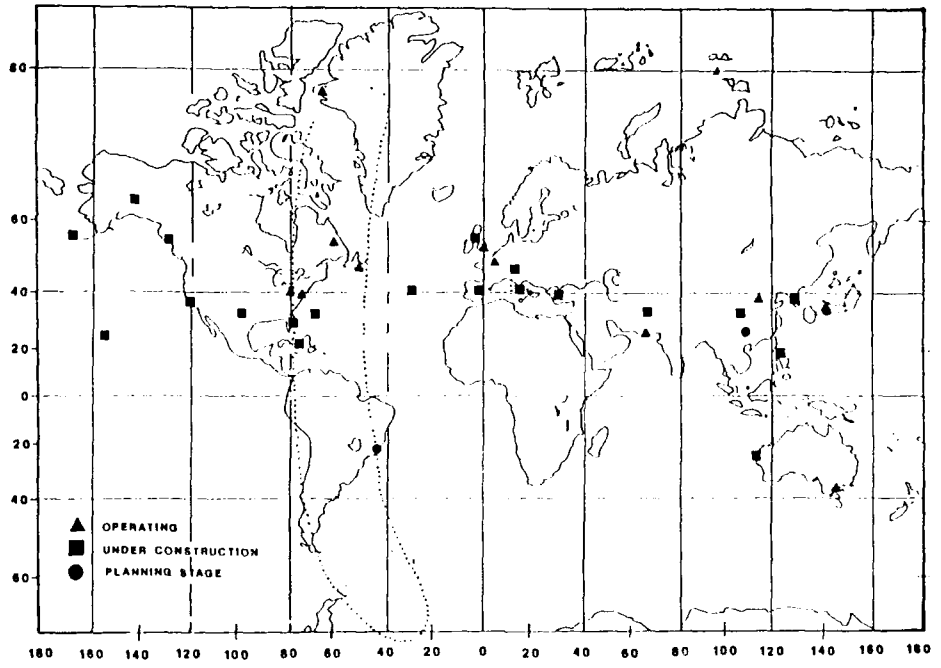
*Nominal Location

Table 1

There are currently 32 Digisonde 256 systems (Table 1) in operation or are close to being installed. The global station distribution, as shown in Figure 1, is very uneven, the majority of sites lying in the northern hemisphere, and there are no equatorial stations. Nevertheless this network provides an extensive data base of ionospheric parameters in digital form, making it easy to process and analyze the data in terms of average diurnal variations, storms, and irregularities. This data base will be invaluable for the testing of global ionosphere models.

The scaled ionogram parameters and/or the raw ionograms can be remotely accessed via voice grade telephone lines operating at 1200 or 2400 baud. Currently only the U.S. Global Air Weather Service is networking in a real time mode by polling all their Digisondes and centrally collecting the ionospheric parameters as soon as an ionogram scan and the 20 to 30 second ARTIST processing is completed. In general, all stations record the raw ionograms and the autoscaled data on half-inch magnetic tape (1600 bpi).

By adding transmit antennas for oblique transmissions to the stations the Digisondes can be used to record oblique digital ionograms (Reinisch et al., 1984, Ahmed et al., 1985). A project is underway to automate both operation and scaling for oblique ionograms, thus adding more sampling points to the global data set.




**GLOBAL DIGISONDE 256 NETWORK**

Figure 1

3. Ionospheric Data

The Digisonde 256 has three basic modes of operation: (1) vertical incidence ionograms, (2) drift observations, and (3) oblique incidence (bistatic) ionograms. Since the third mode is not yet fully automated and the necessary processing software is not yet developed, we limit the discussion to modes 1 and 2.

3.1 Ionogram Characteristics

The Digisonde scales the ionograms within 20 to 30 seconds after completion of the ionogram scan using the ARTIST routine (Reinisch and Huang, 1983). Figure 2 is an example of a typical on-line printout from the University of Lowell station at Millstone Hill in Westford, MA (42.6°N, 71.5°W geogr.). The small numbers using optically weighted font Patenaude et al., 1973) give the amplitudes in multiples of 4 dB for the vertical incidence echoes with ordinary polarization, the X indicates extraordinary polarization, and the arrows point to the direction from where oblique echoes are received. The arrows composing the Es trace at 130 km all point to the north-east. The oblique F echoes between 4.5 and 6.5 MHz come from the south and southeast. ARTIST finds the leading edge of the overhead echo traces for the E and F layers. The autoscaled h'(f) traces, marked by the letters E and F, are superimposed on the raw ionogram, thus providing a means of checking the ARTIST performance. The letter T is used for presenting the calculated true height profile. For each layer, F, F1 and F2, the profile is given in terms of a modified sum of Chebyshev polynomials (Gamache et al., 1985). The coefficients of the Chebyshev polynomials are stored on magnetic tape and transmitted to the data center via telephone links. ARTIST scales the following ionogram parameters: foF2, foF1, foE, foEs, MUFF2(3000), fminF, fminE, h'F, h'F2, h'E and h'Es. The frequency spread for both E and F is determined, and also the average range spread (listed as FF, FF, QE and QF in Figure 2). The virtual height traces h'(f)E and h'(f)F are recorded together with the measured echo amplitudes and Doppler frequencies.

Is the quality of the autoscaled data adequate for the high frequency radio communication and radar engineers, and for the ionospheric modelers? Recent comparisons of ARTIST scalings for foF2 and MUF(3000) with manual readings by experienced ionogram scalers (Gilbert and Smith, 1988) show that ARTIST provides acceptable values for about 93% of the time at a mid-latitude station. For 97% of the analyzed ionograms, foF2 was determined within ±0.5 MHz. This is somewhat better than the 87% (Reinisch and Huang, 1983) found for the auroral station at Goose Bay (64.6°N CGL).

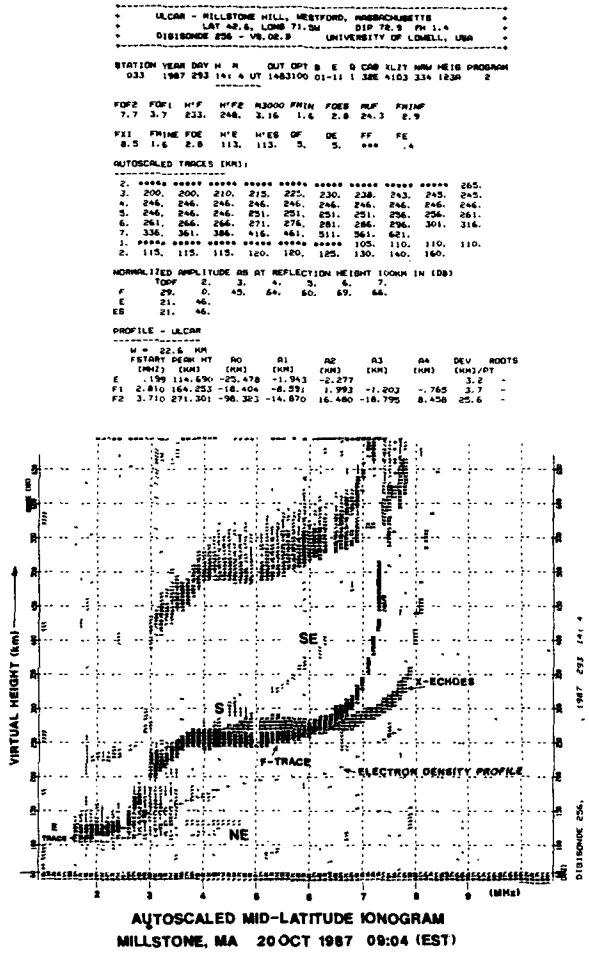



Figure 2

The ULCAR group has recently developed the ARTIST Data Editing Program (ADEP) to establish a quality control for the ARTIST data. Inconsistent values are automatically detected, the raw ionogram together with the ARTIST traces are displayed on a computer display screen, and an operator can make the required corrections. The electron density profiles are updated automatically. The edited ARTIST data are then recorded on a 1.2 Mbyte floppy disk which can store the characteristics for about 1,200 ionograms. This floppy disk will be the form in which the data of the ULCAR station will be published, containing much more information than is normally published by ionospheric stations. Besides the standard ionospheric parameters the disk contains the echo traces, i.e. h'(f), the echo amplitudes and Doppler frequencies and the coefficients for the calculation of the electron density profiles.

A companion program to ADEP prints frequency plots and electron density contours, as shown in Figure 3, displaying the data for 20 October 1987 at Millstone. foE, foEs, foF1 and foF2 are plotted in the top panel. The extent of the frequency spread of the

F trace is indicated on the foF2 curve. The lower half of the figure shows the height variations for given electron densities (plotted as plasma frequency contours in 0.5 MHz increments); the top curve is hmF2. Local midnight and noon are marked by M and N, and sunrise and sunset is marked by an F and an E for the F and E region, respectively. It is proposed that this display become the standard presentation for the data from the Digisonde network. For specific case studies it may be useful to review the original ionograms. The ionogram survey (Figure 4) displays the 96 quarter hour ionograms for one day on a single page. Only overhead echoes with ordinary polarization are printed, oblique and X polarization echoes are deleted.

3.2 Drift Mode Data

In the drift mode, the Digisonde determines high resolution Doppler spectra for each of the signals received on seven spaced receiving antennas (Reinisch et al., 1987, Buchau et al., 1987). The incidence angle of each spectral signal component is calculated from the seven respective phases, resulting in a skymap that shows the location of the simultaneously existing reflection points (sources) and their respective Doppler frequencies. Tight clustering of the sources indicates a smooth ionosphere, wide spread a rough ionosphere. We defined a roughness index RI that is proportional to the angular extent of the source region. Figure 5 shows this index for Erie, CO, during a disturbed period in March and a quiet period in April. The solar zenith angle is plotted on tope of the figure. The roughness index of 2 at noon indicates an angular size of the reflection area of about 10°, while RI = 12 at night during the disturbed period indicates 60°. This pattern of smooth daytime and rough nighttime contours was seen on all the Colorado data and also on data from the subauroral station at Argentia, Newfoundland (Figure 6). The night RI increases with increasing magnetic activity as illustrated in Figures 5 and 6.

This center of gravity of the source locations defines the tilt vector consisting of a tilt angle, measured from the overhead position, and tilt direction, measured clockwise from true north. The diurnal variation of these two parameters for Erie, CO is shown in Figure 7 for 5/6 March 1985. These tilt parameters are well defined only when the ionospheric roughness is small. There is a large uncertainty in defining the tilt parameters when the sources are spread over the sky. Figure 7 indicates a consistent southward tilt increasing from 2° to 3° during the daytime. Large angles of up to 10° are occasionally observed.

When the sources are sufficiently spread the three-dimensional drift velocity vector which best reproduces the observed Doppler frequencies (Dozois, 1967) is calculated. In general, we assume a uniform plasma velocity over the entire skymap (zenith angle $\theta < 40°$). Routine observations in the polar cap and the auroral regions (Buchau et al., 1987) show the predominantly antisunward plasma convection for long periods of time.

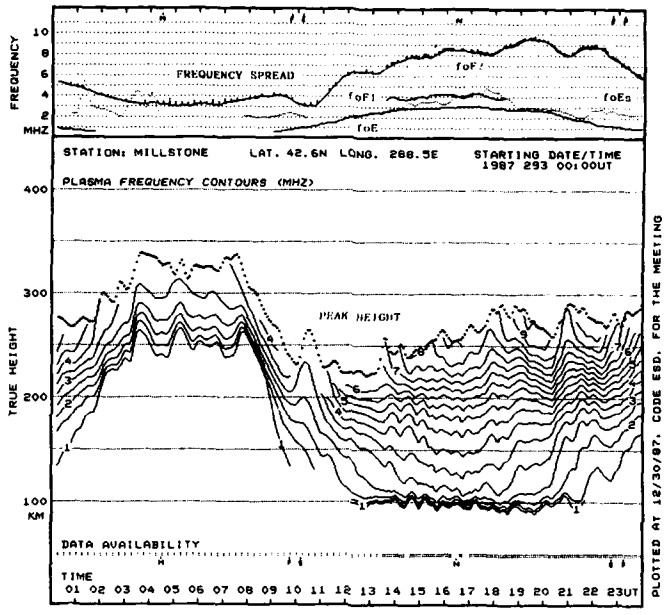
The F region drift in the central polar cap (Qanaq, Greenland, 87.6°N CGL) is shown for five consecutive days, October 18 to 22, 1987 in Figure 8. Direction and magnitude of the horizontal drift vector are plotted against universal time. For a southward z component of the interplanetary magnetic field (IMF) one expects an antisunward drift (Cauffman et al., 1972). Since the IMF data were not available for the period covered in Figure 8 we listed the Kp values. As seen here and also in other data, the drift direction is generally antisunward when Kp > 2 while strong deviations in the sunward direction are observed when Kp is very small. In the period covered by Figure 8, the most consistent antisunward convection occurs from 13 UT on 20 October to 21 UT on the next day, when Kp varies between 2 and 4. A 90° deviation is observed from midnight to 12 UT on 20 October. October 22 shows sunward convection for about 8 hours starting at 12 UT when Kp = 1. The Qanaq ionogram survey (vertical O trace only) for 20 October (Figure 9) reveals that the F region was disturbed until 12 UT, and the F traces show spreading and forking. After 13 UT, when the drift is consistently antisunward, the ionogram signatures are different, suggesting a fairly smooth ionosphere, even though Kp increases to 4.

4. Conclusions

The growing Digisonde 256 network makes it possible for the first time to obtain electron density profiles of the bottomside ionosphere on a global basis with good time resolution. All the data reduction is done in real time leaving only data editing and display to be done off line. It seems that this data base can provide the test bed for the ionospheric modeling efforts. A world-wide study of the ionospheric roughness and the ionospheric drift could give new inputs to the modelers. It will be important to conduct coordinated measuring campaigns to study large scale phenomena like atmospheric gravity waves.

Acknowledgement

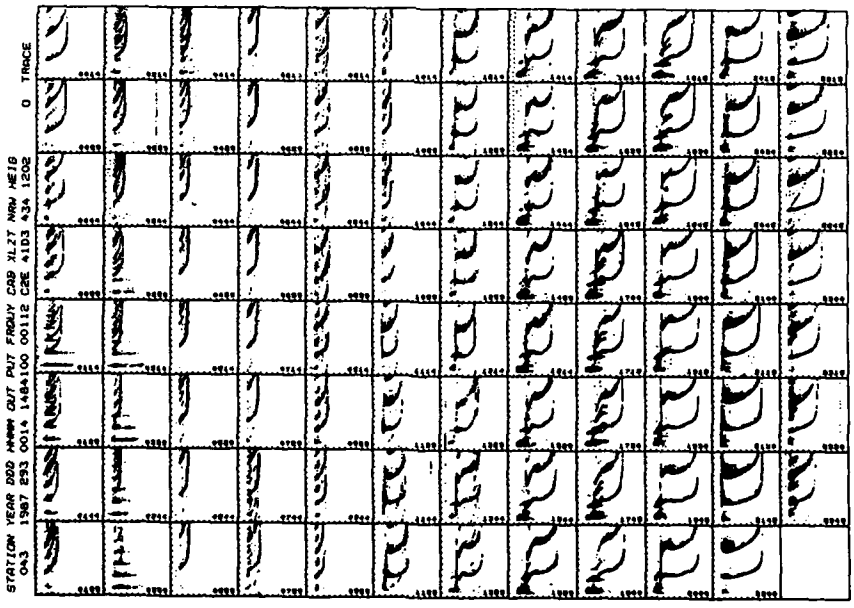
This work was in part supported by the University of Lowell and in part by the Air Force Geophysics Laboratory, Hanscom AFB, Massachusetts.




**IONOSPHERIC CHARACTERISTICS**

**MILLSTONE, MA 20 OCT 987**

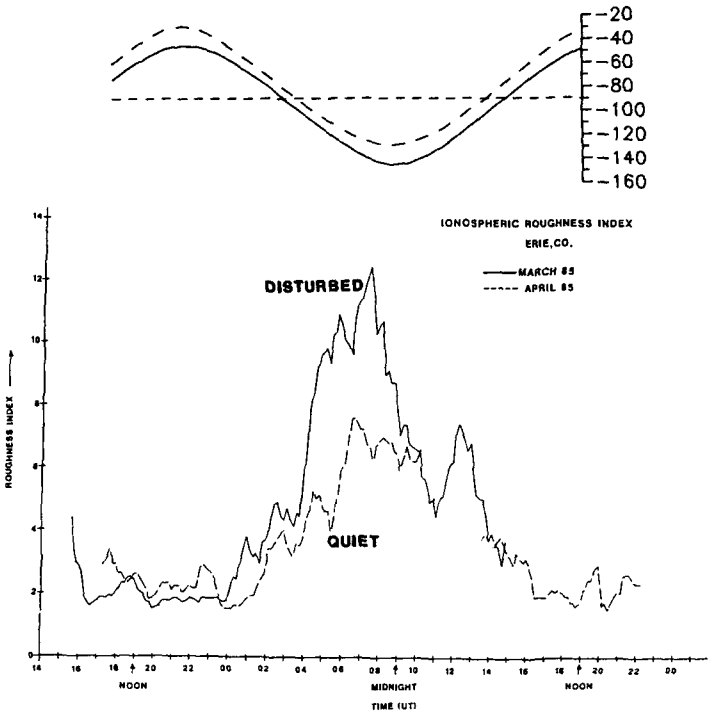
Figure 3




**IONOGRAM SURVEY (X Trace and Oblique Signals Suppressed)**

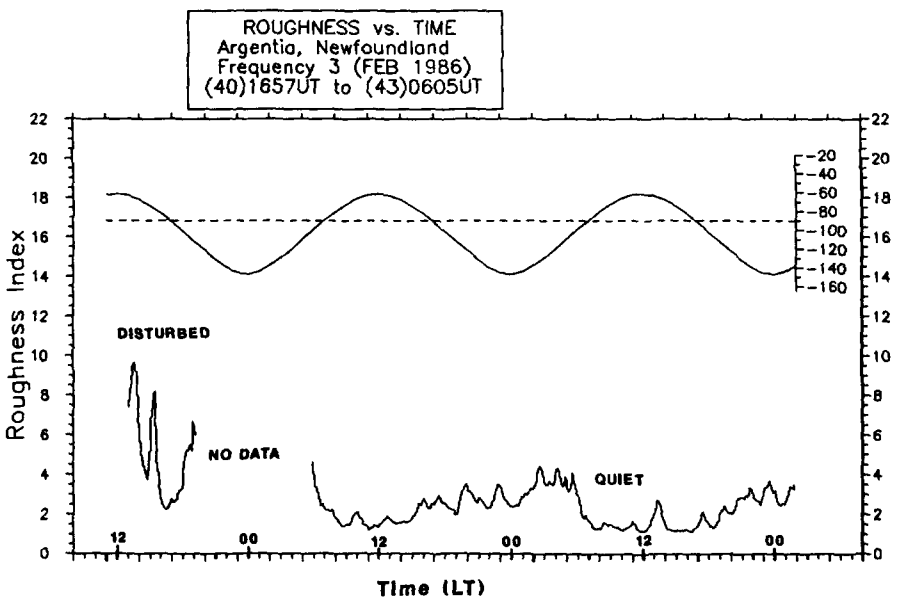
**MILLSTONE, MA 20 OCT 1987**

Figure 4



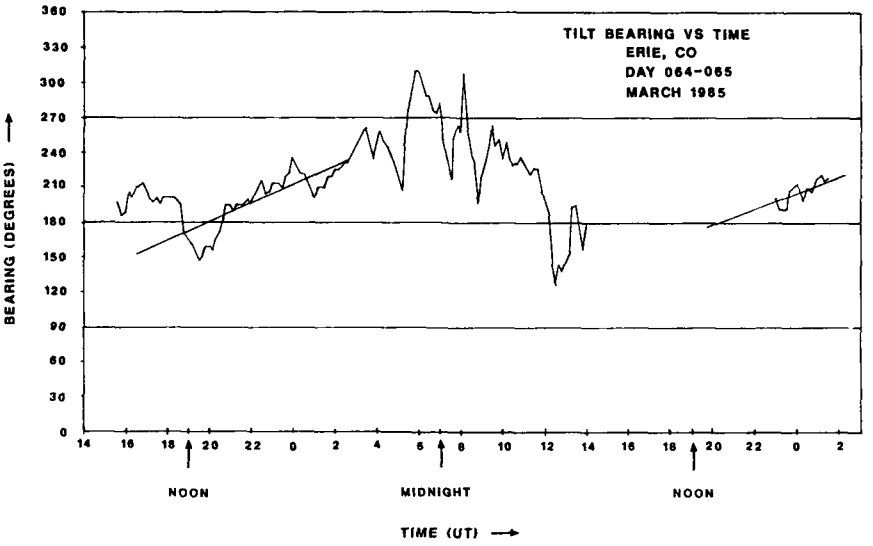

DIURNAL VARIATION OF ROUGHNESS INDEX

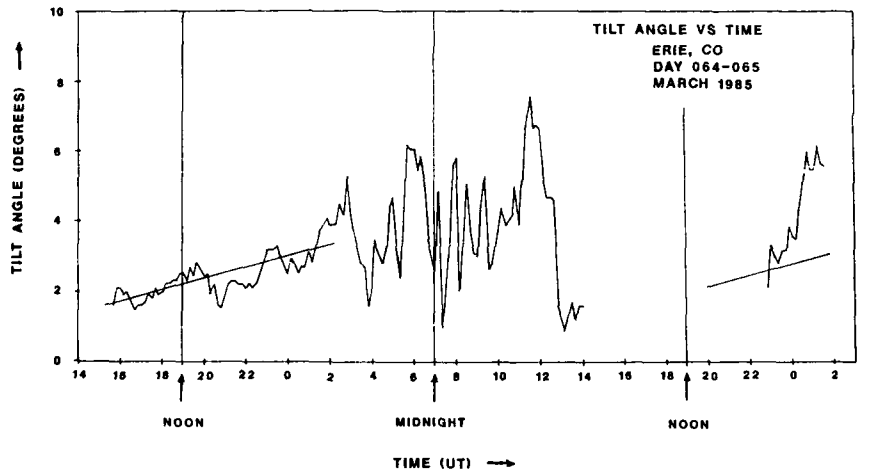
Figure 5




IONOSPHERIC ROUGHNESS INDEX FOR SUBAURORAL F LAYER

Figure 6




TILT ANGLES FOR MID-LATITUDE F LAYER

Figure 7

F-REGION DRIFT
DIGISONDE OBSERVATIONS
QANAQ, GREENLAND

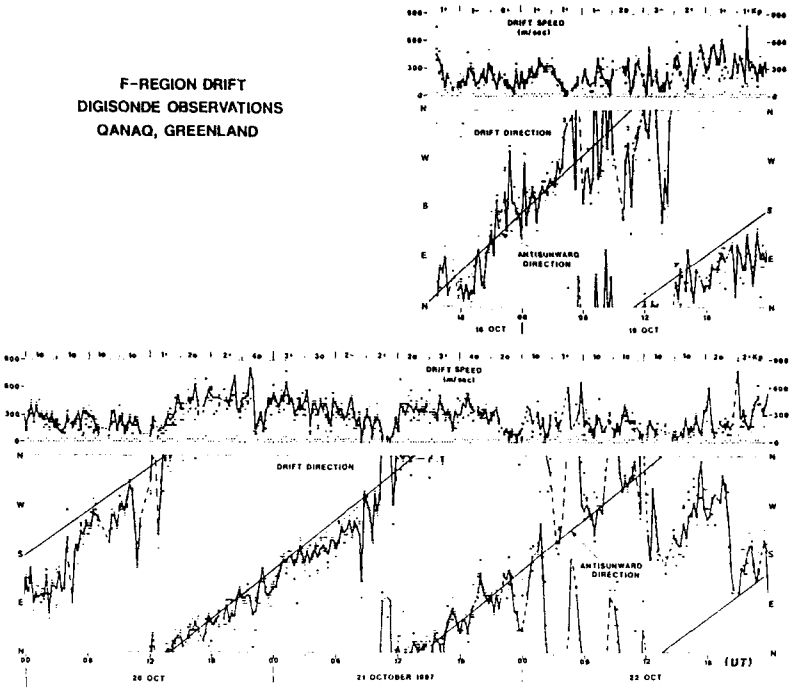
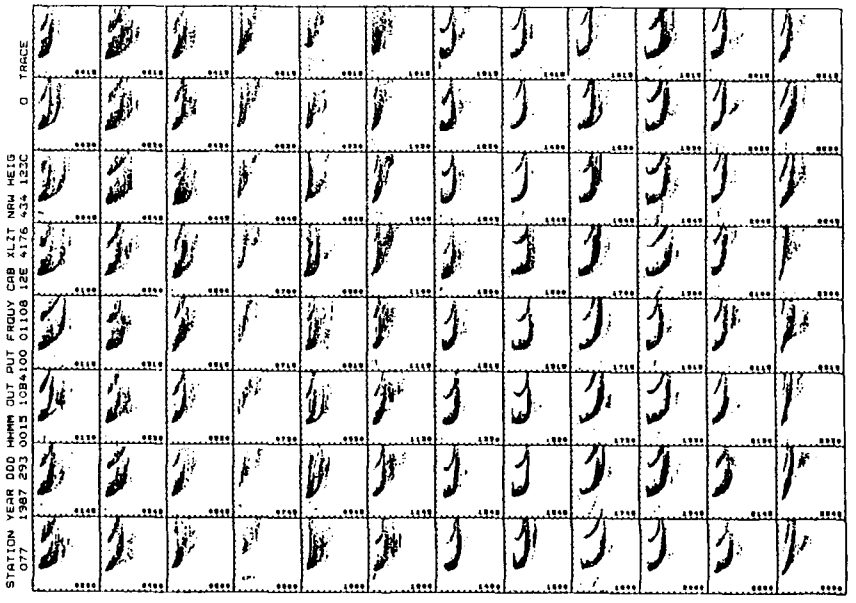



Figure 8




IONOGRAM SURVEY (X Trace and Oblique Signals Suppressed)
QANAQ, GREENLAND 20 OCT 1987

Figure 9

References

Ahmed, M., G. S. Sales and B. W. Reinisch, "Frequency Management of a Long Range HF Communication Link US-UK Observational Data," Proceedings of the IEEE MILCOM Conference, pp. 289-292, October 20-23, 1985.

Buchau, J., B. W. Reinisch, D. N. Anderson, E. J. Weber and C. Dozois, "Polar Cap Plasma Convection Measurements and Their Relevance to the Real-Time Modeling of the High Latitude Ionosphere," Proceedings of IES 87 Symposium on the Effect of the Ionosphere on Communication, Navigation, and Surveillance Systems, May 5-7, 1987.

Dozois, C. G., "A High Frequency Radio Technique for Measuring Plasma Drifts in the Ionosphere," Scientific Report No. 6, AFGL-TR-83-0202, Hanscom AFB, Massachusetts, July 1983.

Gamache, R. R., W. T. Kersey and B. W. Reinisch, "Electron Density Profiles from Automatically Scaled Digital Ionograms. The ARTIST's Valley Solution," Scientific Report No. 1, AFGL-TR-85-0181, 1985.

Gilbert, J. D. and R. W. Smith, "A Comparison Between the Automatic Ionogram Scaling System 'ARTIST' and the Standard Manual Method," submitted to Radio Science for publication in 1988.

Patenaude, J. A., K. Bibl and B. W. Reinisch, "Direct Digital Graphics, The Display of Large Data Fields," American Laboratory, pp. 95-101, September 1973.

Reinisch, B. W., "New Techniques in Ground-Based Ionospheric Sounding and Studies," Radio Science, 21, No. 3, pp. 331-341, May-June 1986.

Reinisch, B. W. and Huang Xueqin, "Automatic Calculation of Electron Density Profiles from Digital Ionograms. 3. Processing of Bottomside Ionograms," Radio Science, 18, 3, pp. 477-492, May-June 1983.

Reinisch, B. W., M. Ahmed, K. Bibl, H. Soicher, F. Gorman, J. D. Jodogne, L. Bossy, J. King and J. Gilbert, "A Transatlantic Digital HF Radio Link Experiment," Proceedings of the Ionospheric Effects Symposium, Effect of the Ionosphere on $C^3I$ Systems, Alexandria, Virginia, pp. 111-122, May 1-3, 1984.

Reinisch, B. W., J. Buchau and E. J. Weber, "Digital Ionosonde Observations of the Polar Cap F Region Convection," Physica Scripta, 36, pp. 372-377, 1987.

## DISCUSSION

**R.Showen**

How are the velocities determined? At what frequencies are they measured with respect to the critical frequencies? Are they radial or inferred horizontal velocities?

**Author's Reply**

The velocity is determined like for a multi beam radar. The transmit antenna illuminates a large area of the ionosphere. Wherever the perpendicularly condition is satisfied, the reflected signals return to the receiving antenna array. The plasma motion $\mathbf{V}$ imposes a Doppler on each reflected signal that is $d = \frac{1}{\lambda}\mathbf{k} \cdot \mathbf{V}$. Discrete Fourier transform in the Dijisonde separates the different reflection points by spectral analysis. The locations of the reflection points are determined by the receiving array ($\sim$ 100 m baselines) using the phases for each Doppler line. A three-dimensional velocity vector $\mathbf{V}$ is then fit by the method of least squares errors to the source location/Doppler field.

**K.Rawer**

Conserving past data is important in geophysics and for application as well. How do the stations archive their data? What kind of compressed information is going to the (National or World) Data Center?

**Author's Reply**

The majority of Dijisonde stations record the digital ionograms and the scaled data on digital magnetic tape (1600 bpi; 4 blocks of 4096 bytes for each ionogram.) The tapes from AFGL stations are stored at AFGL and ULCAR for several years, and then they are transferred to WDCH in Boulder. The Millstone data are now edited with ULCAR's Automatic Data Editing Program (ADEP) and the edited data (h'(f) traces with Doppler, spread and amplitude, the parameters, and the profile coefficients) are written on floppy disk with data for 1200 ionograms per disk. These disks are sent to WDCA.

# COORDINATED MULTI-PARAMETER MEASUREMENTS AND PREDICTIONS OF THE GLOBAL-SCALE IONOSPHERE

Edward P. Szuszczewicz
Laboratory for Atmospheric and Space Sciences
Science Applications International Corp.
McLean, Virginia 22102

R. A. Wolf
Department of Space Physics and Astronomy, Rice University
Houston, Texas 77251

B. G. Fejer and R. W. Schunk
Center for Atmospheric and Space Science, Utah State University
Logan, Utah 84322

E. Roelof
Johns Hopkins Applied Physics Laboratory
Laurel, Maryland 20707

## SUMMARY

The solar-terrestrial system involves a myriad of macro- and microphysical processes that challenge the development of a unified physical description of solar, magnetospheric, interplanetary, and thermospheric controls of the global-scale ionosphere. A well-planned interdisciplinary effort combining theoretical modelling with a coordinated measurement program should lead to significant advances in our understanding of the solar-terrestrial system as a whole and the development of a real-time predictive capability for a world-wide specification of ionospheric conditions as they are dictated by the controlling subsystem inputs.

The SUNDIAL program has such a focus, with a combined modelling and measurement program that includes solar, solar wind, interplanetary and geomagnetic data, and nearly 70 ionospheric monitoring stations at high-, middle- and low-latitudes in the American, European/African, and Asian/Australian sectors. The program operates in an 8-14 day campaign mode with around-the-clock measurements. Campaigns were initiated in October 1984, and since September 1986 have been conducted at 9-month intervals to study each of the four ionospheric seasons twice within the ascending phase of the current solar cycle. We review the measurement and modelling elements in the the program, discuss perspectives on global-scale ionospheric predictability, and highlight some of the accumulating results which address magnetospheric substorms and the triggering of equatorial spread-F.

## 1. INTRODUCTION

To understand coupling processes in the system of solar-terrestrial plasmas requires the synthesis of solar, interplanetary, magnetospheric, thermospheric and ionospheric physics. From certain perspectives, the ionosphere can be viewed as the culmination of all the contributions, and the accurate specification of ionospheric electron density distributions at any time and under any conditions can be viewed as a critical test of our understanding of the myriad of coupling terms. This is the perspective of the SUNDIAL investigation [1] with its objective being the development of a comprehensive understanding and predictive capability for the causal relationships which control the quiescent and disturbed global-scale ionosphere. While the focus is the ionospheric state at any place, at any time and under any condition, the program requires and measures solar and interplanetary inputs, and works to establish a detailed understanding of the interactive roles that the ionosphere plays with the magnetosphere and thermosphere.

Aspects of ionospheric specification and prediction involve: (1) its "quiet time" altitude profile at any place, at any time, and for any season and period within the sunspot cycle; (2) its global-scale response to dynamic events; and (3) its irregularity scale-size distributions under unstable geoplasma conditions. The time-frame classification of predictions is another important aspect, because it helps provide a qualification on the type of prediction and helps focus attention on the fundamental input requirements.

Current advances in solar-terrestrial physics suggest that accurate ionospheric prediction is achievable at two temporal levels. First is the intermediate term, involving periods of tens of minutes, hours, and days, with the upper limit being the 27-day solar rotation period. It is in the intermediate term that the best possibility exists for the development of genuine ionospheric prediction capabilities. In the longer term, periods of months and years, the best that can be expected is a prediction of "average" ionospheric behavior (i.e. ionospheric climatology) for input terms characteristic of the solar and seasonal period for which the prediction is developed. Clearly, a systematically-improved intermediate-term prediction capability will enhance the quality of the predictions for the longer term.

Omitted from the above discussion have been considerations given to short term predictions involving times of the order of seconds to tens of minutes. Accumulated knowledge suggests that predictions in this time frame are largely faulted by nondeterministic processes and that the loosely scientific approach of persistence is the only tool currently available.

## 2. THE MULTI-MEASUREMENT AND MODELLING APPROACH OF SUNDIAL

### 2.1 The Ground-Based Program and Supporting Satellite Activities

The SUNDIAL campaigns have made several contributions to an increased understanding of solar-terrestrial coupling processes and an improved capability in intermediate and long-term predictions [1-7]. In the development of the program elements (see *e.g.* Figure 1), a coupling of theory with experiment was established as an "a priori" condition for tests of our developing understanding. Concentrated and coordinated periods of global-scale measurements and intimate involvement of the site investigators were also established as a definitive means for understanding ionospheric responses, associated coupling processes, and determining a quality data base for comparison with theoretical predictions and existing empirical models.

The ionospheric monitoring network that supports the SUNDIAL effort involves nearly 70 stations covering high-, middle-, and low-latitudes in the American, European/African and Asian/Australian sectors (see Figure 2 and Table 1). The program operates with an agreement on common data formats and works to accomplish around-the-clock data collection on a minimum-resolution quarter-hourly basis for campaign periods 8-14 days in length. Solar, solar-wind, interplanetary, and geomagnetic data are also included through the NOAA/SESC World Data Center.

The ground-based ionospheric measurement techniques used in the SUNDIAL investigation include ionosondes, backscatter radars polarimeters, magnetometers, scintillation receivers, all-sky and scanning photometers, and Fabry-Perot interferometers, with some site-specific capabilities presented in Table 1.

The SUNDIAL ground-based swept-frequency ionosonde network represents a most-valuable cost-effective data acquisition resource. Measurements are carried out with maximum temporal resolution (approximately one-to-five minutes for a complete trace) at a minimum duty cycle of once every 15 minutes (with some stations operating routinely on a 5 minute duty cycle) for the entire 8-14 day campaign period. The processed ionosonde data, reported in an agreed-upon uniform format, represent daily time plots of the peak F-region density (through $f_0F_2$), and three relevant heights ... an estimate of the F-peak ($h_pF_2$, the height at 80% $f_0F_2$), the 3 MHz plasma frequency ($h'_3$), and the lowest edge of the bottomside ($h'F$). In addition, any development of range spreading is represented by a numerical code, with numbers 1, 2, and 3 depicting at any frequency the degree of spreading in the ionogram for $\leq$ 100 km, 100-200 km, and $>$ 200 km, respectively. Furthermore, E-region characterizations (represented by $f_tE_s$ and $f_bE_s$) are included in the total data presentation and real-height profiles are generated using the POLAN program of J. E. Titheridge.

For the ionosonde network the SUNDIAL scientists have developed a PC-compatible data acquisition system called CIDDAS (Computerized Ionosonde Data Display and Acquisition System). Participating stations enter their data on CIDDAS floppies, which include uniform plot formats and provide a relatively inexpensive vehicle for data archiving, distribution, and large volume processing when uploaded into a main frame.

The overall ground-based diagnostics system (ionosondes, radars, etc.) is routinely complemented with relevant satellite observations. To date, major contributions have included measurements from IMP-8, NOAA, TIROS and VIKING.

### 2.2 The Modelling Program and It's Synergism with Observations

An overall perspective on requirements. The overall goal of the program is the development of a predictive capability for the condition of the ionosphere in its quiescent and disturbed state. In reducing this statement to its most fundamental definition, there arises a requirement for the prediction of the ionospheric plasma density profile $N_e(z)$ at all latitudes and longitudes, and for any time. To properly predict $N_e(r,\theta,\phi,t)$ one must necessarily take account of all contributing processes involving production, transport, and loss.

Accurate specification of $N_e(r,t)$ is the ultimate objective of SUNDIAL code prediction, and accordingly is shown as the final output in the simple flow diagram of the measurement and modelling activities in the program (see Figure 3). The zeroth-order input requirements (left-most box in Figure 3) trace themselves to the Sun (UV radiation is implicit), but require only definition of the interplanetary parameters involving the solar wind density and velocity, and the interplanetary magnetic field and its vector direction. Coupling these primary input terms to the required output are two first principle codes (the Rice Convection Model and the Utah State Global Ionospheric Model, discussed in detail in References 1, 2, and 3) which treat the

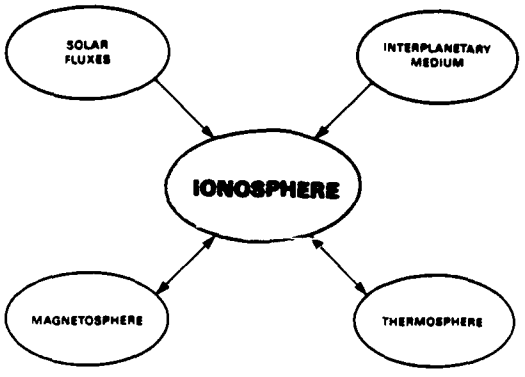



Figure 1. Solar-terrestrial elements in the SUNDIAL program.

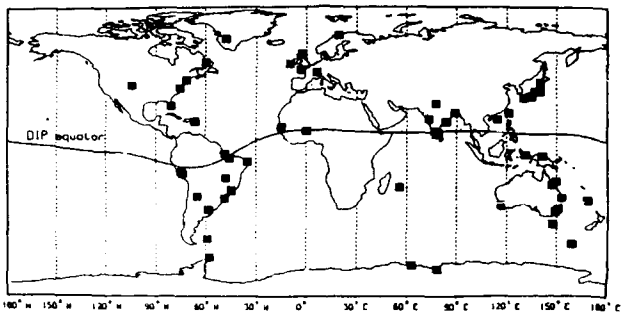



Figure 2. SUNDIAL network of ionospheric stations.

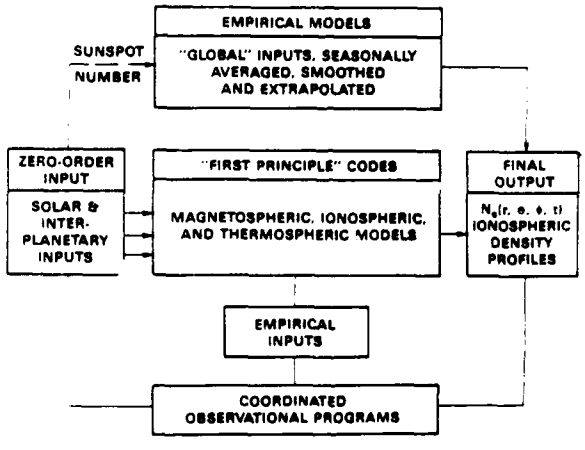



Figure 3. The modelling and measurement approach of the SUNDIAL program.

## TABLE 1

A subset of SUNDIAL ground stations and major facilities.

| REGION | SITE | COORDINATES (degrees) | FACILITIES* |
|---|---|---|---|
| Africa | Dakar | 15 N, 15 W | 1 |
| | Ouagadougou (Upper Volta) | 5 N, 0 W | 1 |
| | La Reunion | 21 11' S, 55 35' E | |
| Argentina | Buenos Aires | 34 5 S, 58 30' W | 1 |
| | Decepcion Is. | 63 S, 58 W | 1 |
| | Falkland Is. | 51 50' S, 59 W | 1 |
| | Tucuman | 26 54' S, 65 24' W | 1 |
| Antarctica | Mawson | 67 60' S, 62 90' E | 1 |
| | Davis | 70 S | |
| Australia-Tasmania | Brisbane | 27.50 , 152.90 E | 1 |
| | Canmen (Syd. ey) | 34 S, 150 40 E | 1 |
| | Canberra | 35.30 S, 149 E | 1 |
| | Hobart | 42 90' S, 147 30' E | 1 |
| | Macquarie Is. | 54 50' S, 159 | 1 |
| | Mundaring | 31.99 S, 116.31 E | 1 |
| | Norfolk Is. | 29 S, 168 E | 1 |
| | Townesville | 19.63 S, 146.85 E | 1 |
| Brazil | Belem | 1 25' S, 48 25' W | 1 |
| | Blumenau | 28 S, 49 W | 1 |
| | Brasilia | 15 52' S, 48 08' W | 1 |
| | S. J. dos Campos | 23 41' S, 45 W | 3 5 |
| | Fortaleza | 3 52' S, 45 51' W | 1 |
| | Natal | 5 51' S, 35 14' W | 5 |
| | C. Paulista | 22 41' S, 45 W | 1 3 5 |
| Canada | Goose Bay | 53 18' N, 60 20' W | 1 2 3 5 |
| Europe | Graz, Austria | 47 N, 6 E | 1 |
| | Lannion | 48 45' N, 3 27' W | 1 |
| | Lindau/Harz | 52 N, 10 W | 3 |
| | Tromso (EISCAT) | 69 N, 19 E, 78 | 2 |
| | Wick (UK) | 58 5 N, 3 W | 1 2 |
| Greenland | Sondrestrom | 67 N, 48 W | 2 |
| India | Bombay | 19 N, 72.8 E | 3 |
| | Calcutta | 23 N, 88.6 E | 3 |
| | Delhi | 28.6 N, 77 1 E | 5 |
| | Kodaikanal | 10.2 N, 77.5 E | 1 |
| | Ooty | 11.9 N, 76.7 E | 5 |
| | Tiruchi | 10.8 N, 78.7 E | 1 |
| | Waltair | 17.7 N, 83.2 E | 3 |
| Indonesia | Hong Kong | 19 N, 114 E | 1 3 5 |
| E. Asia | Taiwan | 23 N, 121 E | 1 3 5 |
| Australia | Yammo | -2.70, 141.3 E | 1 |
| | Darwin | -2, 131 E | 1 |
| Japan | A | 41 N, 140 E | 1 |
| | B | 36 N, 140 E | 1 |
| | C | 36 N, 137 E | 1 |
| | D | 33 N, 135 E | 1 |
| | E | 32 N, 130 E | 1 |
| Peru | Ancon | 75 W, 12 S | 4 5 |
| | Huancayo | 74 W, 13 S | 1 5 |
| | Jicamarca | 75 W, 12 S | 2 |
| South Pacific | Tahiti | 17 44' S, 149 19' W | 1 |
| United States | Arecibo | 18 21' N, 293 15' E | 2 |
| | Boulder, CO | 39 30' N, 104 30' W | 1 |
| | Cape Canaveral | 27 30' N, 81 W | 1 |
| | Millstone Hill | 42 36' N, 288 30' E | 2 |
| | Ramey, P R | 18 N, 67 W | 1 |
| | Wallops Island | 37 57' N, 284 32' E | 1 |

*Facilities: 1 Ionosonde, 2 Radar, 3 Polarimeter, 4 Optical, 5 Scintillators

magnetospheric and ionospheric domains, respectively. The codes are not "first principle" from a purist's perspective, for they require a number of empirical model inputs to couple themselves to processes that go beyond the current capabilities for a full-scale self-consistent "first-principles" approach. While each of these codes have elements that include thermospheric specifications, the NCAR Thermospheric General Circulation Model (TGCM) of Roble and co-workers [8, 9 and references therein] has been the major theoretical thermospheric component in the program since the 1986 campaign.

The Rice Convection Model. The Rice Convection Model (of Wolf and co-workers) provides a consistent picture of the plasma sheet, ring current, the plasmasphere and ionospheric-related parameters (for details and relevant bibliographies see Spiro *et al.*[3]). It covers the lower latitude part of the auroral zone, beginning roughly at the equatorward edge of region-1 Birkland currents, and extends out to about 10 $R_E$ in the equatorial plane. The auroral electric field profiles tend to be more realistic than power-law descriptions and the model provides an improved representation of electric field penetration to low latitudes. Some of the new features in the code (*e.g.*, the thermospheric wind model) are being tested in the SUNDIAL program.

The Utah State Ionospheric Model. The ionospheric model adopted for test and development in the SUNDIAL program is that of Sojka and Schunk (for details and relevant bibliographies see Szuszczewicz *et al.* [1] and Schunk and Szuszczewicz [2]). It is a time-dependent, 3-dimensional, multi-ion ($NO^+$, $O_2^+$, $N_2^+$, $O^+$, $N^+$, $He^+$) model of the global ionosphere at altitudes between 120 and 800 km. The model takes account of the effects of field-aligned diffusion, cross-field electrodynamic drifts both in the equatorial region and at high latitudes, interhemispheric flow, thermospheric winds, polar wind escape, energy-dependent chemical reactions, neutral composition changes, and ion production by EUV radiation and auroral precipitation.

Altitude profiles of the ion and electron temperatures and the six ion densities are obtained by solving the appropriate continuity, momentum, and energy equations. At middle and high latitudes, the equations are solved over the altitude range from 120 to 800 km, with chemical equilibrium at 120 km and a specified plasma escape flux at 800 km being the lower and upper boundary conditions, respectively. At low latitudes, the temperatures and densities are computed along trans-equatorial flux tubes from 120 km in one hemisphere to 120 km in the conjugated hemisphere.

Empirical modelling. The data acquired in all SUNDIAL campaigns will also provide a valuable resource in improving existing empirical models of the global-scale ionosphere. The empirical models are based on a statistical and/or numerical description of the ionosphere in terms of location (geographic or geomagnetic), time (solar zenith angle), solar activity (10.7 cm flux or sunspot number), and season. By definition, empirical ionospheric models are based on accumulated data, including critical frequencies ($f_0E$, $f_0E_s$, $f_0F_1$, and $f_0F_2$), altitudes of peak concentrations ($h_mE, h_mF_1$, and $h_mF_2$), and half thicknesses of the individual layers. The data is compiled typically from ionosondes, topside sounders, and "in situ" satellite and rocket profiles (wherever available). Averaged profiles are then synthesized, and analytical fits developed with the end product representing an intelligent mix of empiricism, physics, extrapolation, and intuition.

There are a number of contemporary schemes for empirical modelling of ionospheric profiles, with a recent review and relevant bibliography included in Reference 2. The International Reference Ionosphere (IRI) has been adapted for comparative purposes in the SUNDIAL investigations.

The IRI attempts to provide global-scale information not only on the macrocale features of electron density, but also on the ion and electron temperatures, and the ion composition. It is a simple, fast-running PC-compatible code that allows a synthesized perspective on UT, LT, solar-cycle, seasonal and magnetic latitude effects. These features make the IRI an attractive model for comparison with SUNDIAL data and our first-principle code results. It is an evolutionary code which makes no pretense about predictive capabilities, with its primary emphasis being an attempt at a global-scale representation of experimental observations. In this regard the IRI makes its limitations clear, particularly with regard to the relative paucity of ion composition results and inadequacies at auroral latitudes. The founders and contributors to the IRI encourage workers in the field to help fill in the gaps, both in the understanding of ionospheric processes, and in the scope of experimental coverage. We expect that the SUNDIAL activities will be most useful in their theoretical and experimental comparisons with the IRI, providing not only an improved specification of the global-scale ionosphere, but also an adaptive model specification which will allow descriptions of global ionospheric responses to such dynamic events as magnetospheric substorms.

## 3. CAMPAIGN RESULTS

### 3.1 Triggering of Equatorial Spread-F by an Interplanetary Event

The 1984 SUNDIAL campaign made several advances that contributed to our understanding of solar-terrestrial coupling processes and an improved capability in intermediate and long-term predictions [1-7]. It was a campaign that involved eight days of around-the-clock measurements of the global-scale ionosphere, complemented by solar, interplanetary and magnetospheric inputs. Particularly valuable to the ionospheric complement and associated magnetometer measurements were solar observations of a cross-equatorial coronal hole, interplanetary particle and field characterizations of a high speed solar wind stream by the IMP-8 spacecraft, specification of exo-ionospheric inputs to high latitudes by NOAA/TIROS satellite measurements, and the synoptic field and energy input terms developed in collaboration with the Millstone Hill facility. The measurement and modelling activities covered both quiet and disturbed conditions, with emphasis on an accurate specification of the quiescent state of the global-scale ionosphere and the dy-

namic effects of two substorms that occured during the campaign. One of the two substorms resulted in the most intense condition of equatorial spread-F ever recorded. Of particular importance to predictions, the 84 campaign was able to trace coupling processes through a spectrum of events that started with the coronal-hole-generated high-speed solar wind stream. Measurements and models accounted for interplanetary magnetic field dynamics, the development and decline of cross-polar cap potentials, the "stirring" of fossil winds at high-latitudes, associated penetration of attendant electric fields to the equatorial domain and their catalytic effects in the triggering of equatorial spread-F. In the paragraphs to follow, we offer the details of the solar-terrestrial observations, their interactive and corroborative components, their causal contributions to the spread-F event, and their relevance in bounding the modelling efforts. A comprehensive reporting of all SUNDIAL-84 results appears in References 1 - 7. Szuszczewicz *et al* [1] provide an overview of the entire activity, along with a synopsis of operations, observations and attendant analyses of all accumulated data. Schunk and Szuszczewicz [2] provide a perspective on the current status of global-scale ionospheric models while References 1 and 3-7 attempt to pull the observational program elements together with detailed analyses...dealing first with solar and interplanetary observations (Szuszczewicz *et al* [1]), solar wind-magnetospheric-ionospheric coupling processes (Spiro *et al* [3]) and the comparisons of global-scale ionospheric data with existing and developing models (Robinson *et al* [4], Leitinger *et al* [5], Wilkinson *et al* [6], and Abdu *et al* [7]).

SUNDIAL-84 covered the 5-13 October 1984 period which was approaching solar minimum in the declining phase of Solar Cycle 21. It was a flare-free interval that included the rise, maximum, and fall in the velocities of a co-rotating high-speed solar wind stream eminating from the equatorward extension of a polar coronal hole. With the exception of the transition to the dynamics associated with the high- speed stream, solar phenomenology was nominally that of a typical solar minimum condition. Solar activity was very low throughout the SUNDIAL period with the Sun spotless for several days. There were no significant flares at soft-x-ray wavelengths (1-8 Å), or notable solar radio noise events during the period; and only a few optical flares were reported with no significant enhancements in energetic particle fluxes measured at geosynchronous satellite altitudes. The trans-equatorial coronal hole and the associated interplanetary particle and magnetic field dynamics provided the causal terms for magnetospheric and ionospheric disturbances observed at the end of the second day in the SUNDIAL-84 campaign [1]. The geomagnetic field was mostly quiet on the 5th and 6th ($A_p = 7$ on the 5th) and throughout the major portion of the SUNDIAL observation period.

Analyses of high-latitude magnetograms [3] supported the existence of quiescent ionospheric and magnetospheric conditions at SUNDIAL times prior to 2000 UT on 6 October. The period 2000-2100 UT marked the onset of a substantial substorm with full development at 2200 UT (600 $\gamma$ maximum excursion). This onset is reflected in the southward turning of the IMF at 2018 UT on 6 October (see Figure 4) and in the *Dst* and *AE* indices plotted for 6-7 October 1984 in Figure 5. Several of the magnetogram sites indicated the beginning of a recovery phase at about 2220 UT and at 2300 UT all stations had recovered. These results are consistent with the IMF data in Figure 4, which show a southward turning of the IMF at 2018 UT on the 6th, with a sudden northward reversal at 2200 UT. The magnetogram results and attendant AE indices (see Reference 3) also suggest that the IMF maintained its northward orientation throughout the rest of the evening on the 6th (filling in the gap in the IMP-8 data). Magnetogram analysis also points to another substorm in the period 0100-0200 UT on the 7th.

A measure of energetic particle precipitation power at high latitudes and the cross polar cap potential throughout the entire SUNDIAL period was made available through the NOAA/TIROS "precipitation activity index" [10]. The results are presented in Figure 6, with an inspection of the data showing a correlation with the events described on 5 through 7 October. The figure also shows for that period a sustained precipitation power greater than 70 GW and a cross-polar cap potential averaging about 75 kV. These levels of power and potential correspond to a precipitation activity index 8 to 9 and an average $K_p$ between $4^0$ and $5^-$. Further inspection of Figure 6 shows the low input power levels on 5 October and through 20 UT on 6 October, and again in the final two days, 12- 13 October, of the campaign.

The "record setting" SUNDIAL-84 spread-*F* and associated 3m irregularity plume events at the Jicamarca Radar Observatory (Figure 7) followed a rapid increase in the vertical *F*-region plasma drift velocity that had its onset at about 1900 LT [1,3]. This vertical drift velocity enhancement, at a universal time of 00 UT on 7 October, is typical of low latitude electric field penetration events during magnetospheric disturbances. Such events in the dusk sector are often associated with sudden increases in magnetospheric convection (see References 3 and 11). Its occurrence, immediately following the normal evening *F*-region dynamo-induced pre-reversal enhancement in the vertical ionization drift (that had an onset at 18 LT (23 UT) on this day, see Figure 8), is believed reponsible for the intense nature of this unprecedented spread-*F* event over Jicamarca (see References 3, 7 and 12). Simultaneous fluctuations in *F*-layer heights and horizontal (H) components in magnetograms were observed at locations in the American, African, Indian, and Southeast Asian sectors of the equatorial zone (see *e.g.*, Figure 8). Evidence of coupling of the different ionospheric height domains was also detected at a number of geographic regions as well as on a global scale (Abdu *et al.*, 1988 [7]).

The campaign provided two distinct opportunities to model and test developing concepts for interplanetary and high-latitude controls over equatorial F-region dynamics. The two events (one of which was

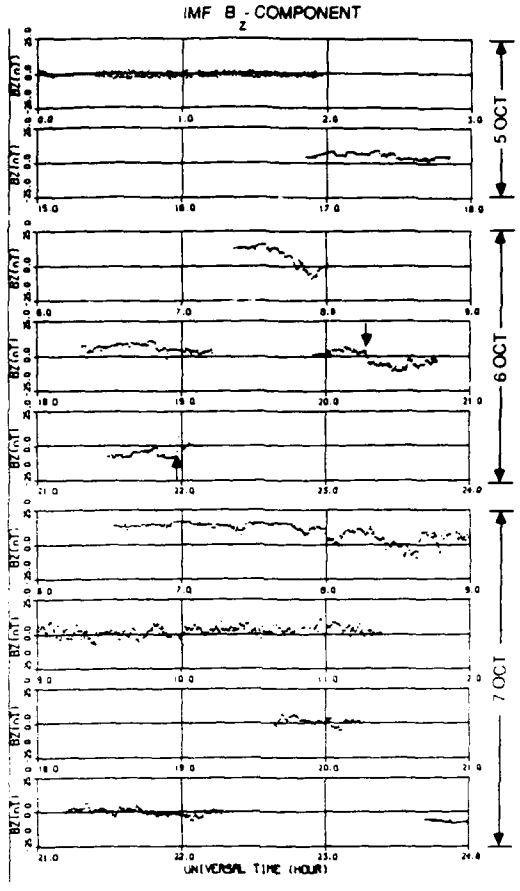



Figure 4. $B_z$-component of the IMF for all measurement periods of the IMP-8 spacecraft on 5, 6 and 7 October 1984. $B_z$ units in nT. (Courtesy of R. Lepping at GSFC).

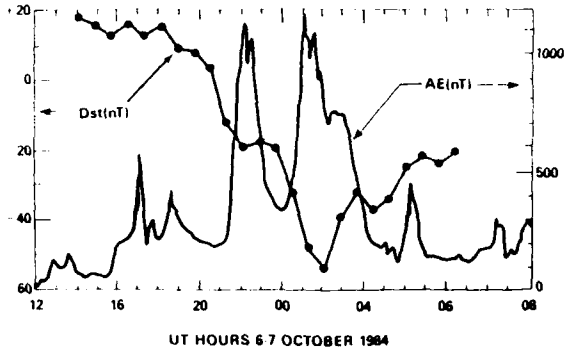



Figure 5. Dst and AE indices for 6-7 October 1984 (Adapted from [3]).

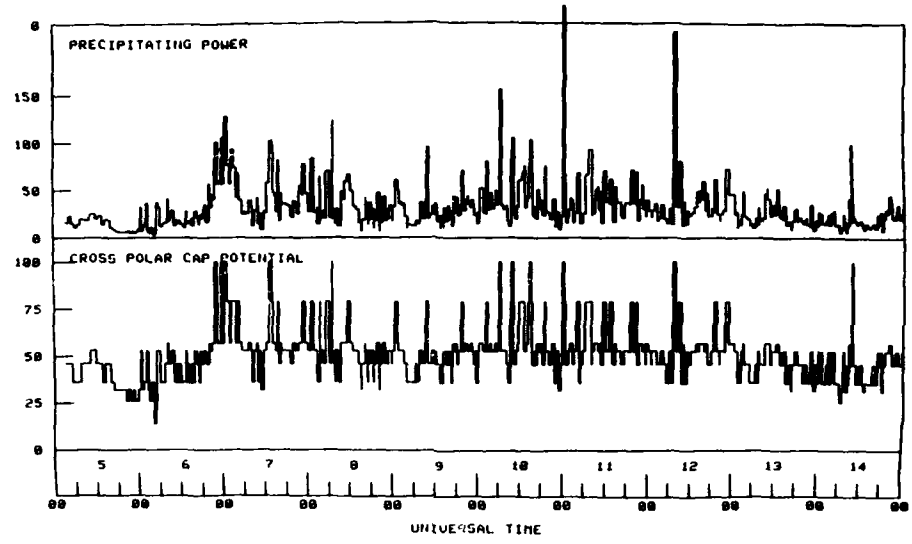



Figure 6. Energetic particle precipitation power at high latitudes and the cross polar-cap potential for the SUNDIAL-84 period as calculated from NOAA/TIROS satellite data (Courtesy of J. Foster, see *e.g. Foster et al.*, 1986).

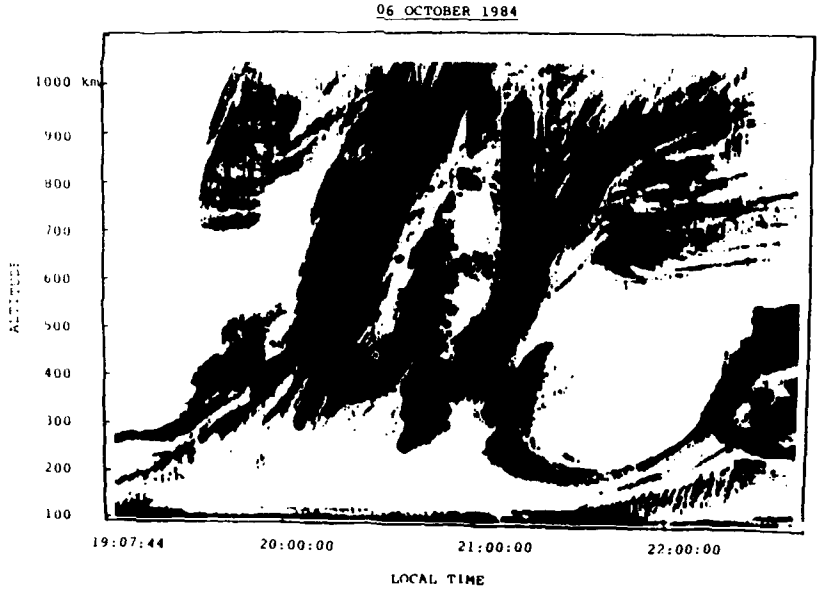



Figure 7 Jicamarca radar plot of reflected energy from domains of 3-m plasma irregularities during the equatorial spread-F event of 6-7 October 1984 (UT=LT + 5hr).

discussed above as "record setting"), represented by the vertical drift velocity measurements in Figures 8 and 9, corresponded to the onset and recovery phases of separate substorms. More specifically, they were associated with southward and northward turnings of the IMF and the associated increase (Figure 8) and decrease (Figure 9) in convection electric fields at high latitudes. Figure 8 presents a 24 hour time base covering 6-7 October 1984, with the smooth solid line representing the typical diurnal trend discussed above. The irregular solid line presents the data available during the late afternoon and early evening hours, while the cross-hatched region from 19 through 3 hr LT depicts an ESF condition which precludes drift velocity measurements. The onset of ESF during the period of pre-reversal enhancement in Figure 8 has been associated with a southward turning of the IMF, an increase in cross polar cap potential from 40 to 90 keV, and an immediate penetration (as measured and modelled) of increased electric fields to the nighttime equatorial domain [3]. The penetrating eastward fields at a time when the equatorial F-region ionosphere was preconditioned for ESF onset is believed to be the catalyst which triggered the "record setting" spread-F event that occurred in the period following 19 hr LT [1, 3]. The numerical modelling experiments which describe the overall coupling of interplanetary plasma and magnetic field parameters through the magnetospheric and high-latitude ionospheric domains to the equatorial F-region are detailed in the work of Spiro *et al.* [3] and will not be further elaborated upon here. Instead, we will discuss the event depicted in Figure 9.

Figure 9 shows that at about 0100 hr LT (0600 UT) the vertical drift velocity increased suddenly by approximately 20 m/s above the normal quiet-time value for that time of night. This change corresponds to an increased eastward electric field of 0.5 mV/m, an enhancement that occurred during the recovery phase of a substorm observed at high latitudes in about the same longitudinal sector [3]. Such a strong eastward electric field perturbation in the midnight-to-dawn sector is a typical observation for a sudden northward turning of the IMF after having been in the southward IMF configuration for one or more hours [11, 13]. To develop insight into the coupling processes and to establish a meaningful quantification for contributing terms, a series of numerical experiments was carried out with the Rice Convection Model (RCM) developed by Wolf and co-workers [3, 14]. The numerical experiments assumed an idealized substorm, patterned after the high latitude time profile of cross-polar-cap potential in the 6-7 October period (see Figure 6). That model supported the interpretation of results in Figure 8 and was employed in attempts to fit the results of Figure 9. Comparison of model results with the observations in Figure 9 led to the following conclusions: 1) Model calculations performed under the assumption of zero neutral wind were consistent with the observations with regard to pattern and maximum strength, but not with regard to duration; 2) The most promising theoretical explanation for the observed 1-2 hr duration of the low-latitude response to decreased convection (see Figure 9) seems to be the "fossil-wind" idea suggested by A. D. Richmond. The second result is of special significance.

The "fossil-wind" concept addressed disturbance electric fields at low latitudes associated with a northward turning of the IMF, suggesting that the low-latitude effects may be traced to the poleward movement of the shielding layer (and the equatorward edge of the auroral oval) when convection decreases. The poleward retreat of the shielding layer can leave behind a high-latitude neutral wind (termed "fossil" wind) with dynamo electric field effects which can penetrate unshielded to low latitudes. The agreement between the model results (with a "fossil wind") and observations helps elucidate the critical importance in the proper tracking of coupling processes that originate with a solar-driven interplanetary disturbance, and couple through the magnetospheric and high latitude ionospheric domains to the equatorial F-region via the forces of auroral oval dynamics, winds and electric fields. While modelling such an event falls in the category of intermediate-term predictability, this does not mean that such a predictive capability is in hand; the results are not completely unambiguous and the fossil wind has only been modelled in an extremely crude way. It does mean that the computer experiments performed with the Rice Convection Model have provided substantial insight into the interpretation of low-latitude ionospheric responses to dynamic interplanetary events, particularly when those events can be classified as discrete and stable turnings of the IMF. This has given credibility to the developing concept referred to as a high-latitude-driven fossil wind, and for the first time focuses on a coupled system that includes the magnetosphere, the ionosphere and components of the thermosphere. While the current RCM model falls short of a self-consistent representation of all terms, it has allowed for an improved understanding of causal components that include tidal winds, F-region dynamo fields and magnetospheric events. Further improvements in the model specification of the magnetospheric, ionospheric and thermospheric domains hold promise for improved accuracy in understanding intermediate term events and attendant extrapolations to improved predictive capabilities.

### 3.2 Follow-On Activities and Adaptive Empirical Models

Since the SUNDIAL-84 activities, three additional campaigns have been carried out covering 8-14 day periods in September-October '86, May-June '87 and March '88. Analysis of results that is furthest along is that identified with September-October '86, to be referred to here as SUNDIAL-86.

As an equinoctial campaign, SUNDIAL-86 represents an excellent follow-up to the '84 effort. Covering the period from 22 September through 4 October, the '86 campaign occurred at the very extreme of solar minimum, with the entire month of September designated a "Solquiet" and the sunspot number at 0 during most of the period. (See synoptic presentation of solar-terrestrial indices in Figure 10 ) For several months before and after the campaign, geomagnetic activity and the solar wind plasma characteristics demonstrated a remarkable (and some would say predictable) 27 day recurrence format. The SUNDIAL period itself began with a very quiet day (22 September), then entered a substantial 3-day storm period late on the 22nd which

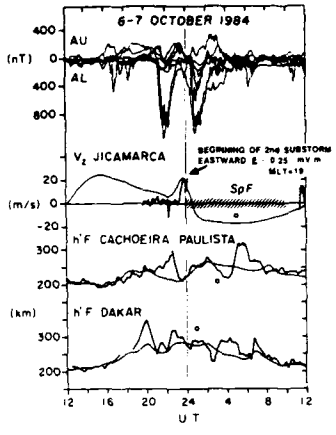



Figure 8. High latitude magnetic field, equatorial F-region vertical plasma drift velocity, and h'F variations during 6-7 October 1984. The dots indicate the times of local midnight.

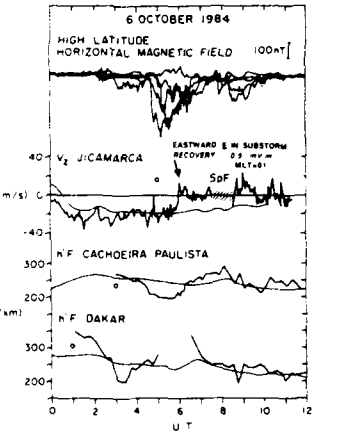



Figure 9. High latitude magnetic field, equatorial F-region vertical plasma drift velocity, and h'F variations during 0-12 UT, 6 October 1984. The dots indicate the times of local midnight.

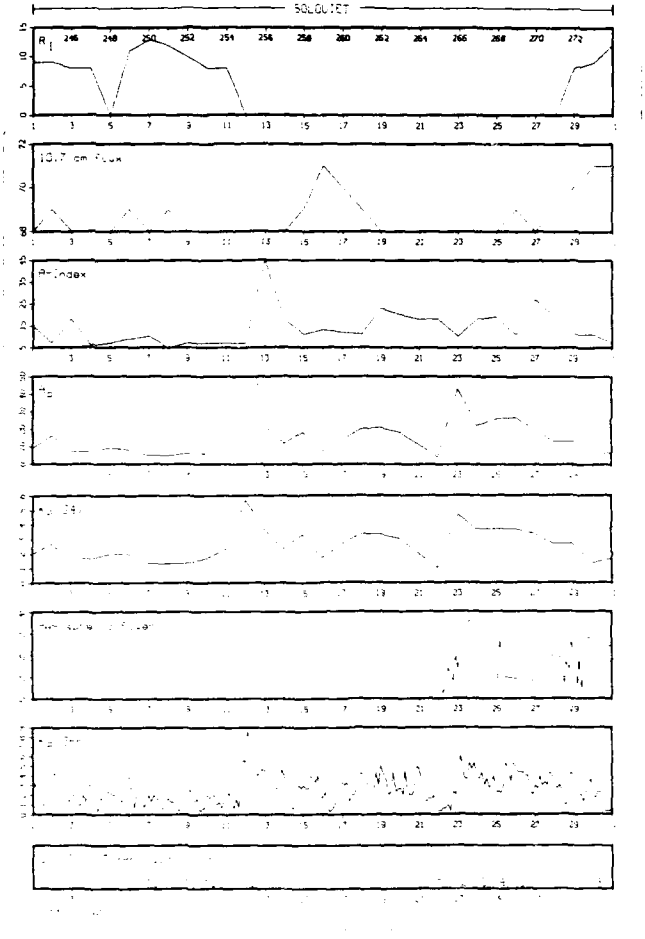



Figure 10. Solar-terrestrial activity indices encompassing the SUNDIAL-86 campaign period. Panels top-to-bottom depict the International Sunspot Number $R_I$, the 10.7 solar flux, the A-index, the planetary $A_p$ index, the 24-hr average $K_p$, NOAA/TIROS hemispheric power, the 3-hr $K_p$ index, and the quiet/disturbed day indicator ($Q_1/D_1$ = most quiet/most disturbed during entire month).

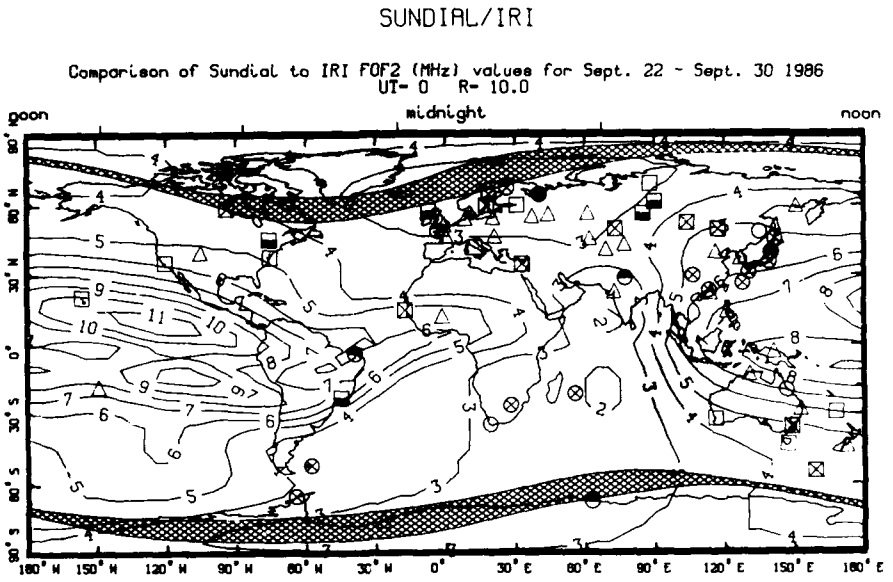



Figure 11. Comparison of $f_0F_2$ (= $N_m F_2$ through plasma frequency relationship) observations with the contour plots of the IRI, including a Feldstein oval specification at $Q = 6$. Δ's specify agreement within ±5%. ⎺, ⊡, ▬, and ■ indicate the observations are less than IRI specifications by 5-10%, 10-20%, 20-40% and >40%, respectively. ○, ⊗, ◓, and ● have the same quantitative scaling but for observations greater than the IRI specifications.

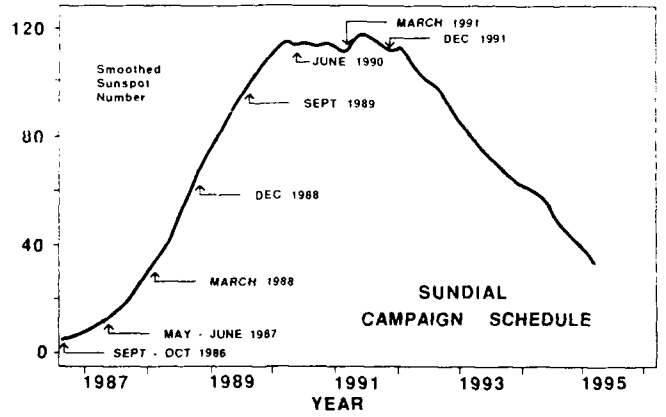



Figure 12. SUNDIAL campaign periods superimposed on predicted sunspot numbers for solar cycle 22.

RISE OF SOLAR CYCLE 22 COMPARED TO PREVIOUS CYCLES
(From NOAA/Boulder)

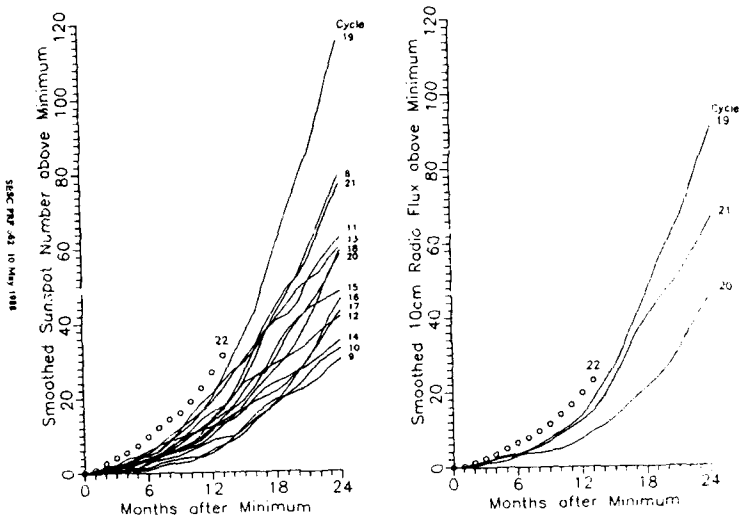



Figure 13.

TABLE 2

PROF[illegible]SF[illegible] PAIGN AND WORKSHOP SCHEDULE

| Campaigns | Workshops |
|---|---|
| December 1988 | |
| | May 1989 |
| September 1989 | |
| | January 1990 |
| June 1990 | |
| | November 1990 |
| March 1991 | |
| | May 1991 |
| December 1991 | |
| | January 1992 |

slowly subsided to another global-scale quiet day on 30 September. The entire period covered the most quiet ($Q_1$) and second most disturbed ($D_2$) days of the entire month of September (see Figure 10).

The results of '86 and developing analyses are providing an excellent opportunity to extend and intensify our understanding of coupling processes in the overall system. Reasons for this include:

i) The long duration campaign itself, which started with a very quiet day, encompassed a major 3-day substorm activity and returned to a quiet condition. The science team is now in a position to analyze and model a global-scale quiet condition and its transition to a progressively disturbed case.

ii) The diagnostics in '86 were superior to those available in '84. SUNDIAL had improved high-latitude coverage with EISCAT, Sondrestrom and Millstone Hill; a much improved complement of thermospheric wind measurement stations; and added satellite coverage provided by VIKING.

iii) SUNDIAL data analysis tools are better developed than those in '84. One such illustration is provided in Figure 11, where we have made use of the CIDDAS data acquisition system in a way that allows VAX computer handling of our global-scale ionospheric observations and model predictions. Figure 11 shows contours of $f_0F_2$ as specified by the IRI for UT=0 under solar minimum conditions (R=10) during the month of September. The symbols represent UT=0 ionosonde-station observations averaged for the September SUNDIAL-86 period. The triangles represent agreement within 5%, while circles/squares represent higher/lower observations than that specified by the local IRI values. As the symbols become filled, disagreement increases from 10% to greater than 40%. The worst case is represented by solid circles or solid squares. Comparisons of this type are now being executed at close intervals of time to study the transition of the global-scale ionosphere from its quiet condition on the 22nd to and through its most disturbed state on the 24th. Current analyses point to very good agreement between hourly 8-day-averaged observations and the global prescription of the IRI. On the other hand, specific time-progressive global plots of observations (*e.g.* starting with 0 UT, 22 September and running through 24 UT, 25 September at 3 hours increments) are revealing effects of auroral dynamics on $N_2/0$ ratios as manifested in reductions in values of $f_0F_2$. The time- progressive plots are also providing the fundamental basis for an adaptive empirically-derived global-scale ionospheric model that will assist in improved accuracy in longer-term predictions.

## 4. PERSPECTIVES AND PLANS

The SUNDIAL objectives are ambitious and the advances made thus far are encouraging. It is recognized however that the near-Earth space plasma environment will change drastically during the rise of this solar cycle. Solar wind disturbances will no longer be due to quasi-steady corotating streams from coronal source regions of weak, open magnetic fields (as in SUNDIAL-84), or to "predictable" 27-day recurrences (as in SUNDIAL-86). Rather, the high velocities will be produced by transient disturbances whose energy is derived from highly dynamic and increasingly active regions. It is with this perspective on increased and prolonged solar dynamics that the SUNDIAL membership has developed the schedule shown in Figure 12 and delineated in Table 2. The campaign schedule and associated workshops will trace the flow of energy and mass in the Sun-Earth system through an increasing sunspot number, unfolding in each campaign and successive workshop the developing complexities in the total system. The schedule provides for a campaign and workshop every nine months, allowing for a systematic sampling of the four ionospheric seasons twice in the transition from solar minimum to solar maximum of cycle 22.

The issue of projected sunspot number for the maximum of cycle 22 is being watched with some excitement, since the current rise rate exceeds even the strongest cycles of record (see Figure 13). There are some expectations that cycle 22 will provide the highest activity of recorded solar-geophysical data. From this perspective, the accumulating SUNDIAL data base and attendant empirical and first principle model improvements will provide a unique opportunity to fulfill objectives, within the widest dynamic range of solar controls ever studied.

## ACKNOWLEDGEMENT

This work has been supported by the National Science Foundation (NSF) under Grant No. ATM-8513362. To the NSF and to all participating countries, institutions and SUNDIAL investigators we extend our sincere thanks for dedicated support.

## REFERENCES

[1] E. P. Szuszczewicz, B. G. Fejer, E. Roelof, R. Schunk, R. Wolf, R. Leitinger, M. Abdu, B. M. Reddy, J. Joselyn, P. Wilkinson and R. Woodman, "SUNDIAL: A World-wide study of interactive ionospheric processes and their roles in the transfer of energy and mass in the Sun-Earth system," Annales Geophys. 6, 3, (1988).

[2] R. W. Schunk and E. P. Szuszczewicz, "First-principle and empirical modelling of the global-scale ionosphere," Annales Geophys. 6, 19, (1988).

[3] R. W. Spiro, R. A. Wolf and B. G. Fejer, "Penetration of high-latitude-electric-field effects of low latitudes during SUNDIAL 1984," Annales Geophys. 6, 39, (1988).

[4] T. R. Robinson, J. A. Waldock, M. D. Burrage and T. B. Jones, "High latitude ionospheric response to changes in the interplanetary medium: SABRE observations during SUNDIAL," Annales Geophys.6, 51, (1988).

[5] R. Leitinger, P. Wilkinson and R. Hanbaba, "The ionosphere in mid-latitudes during the SUNDIAL campaign," Annales Geophys. 6, 59, (1988).

[6] P. J. Wilkinson, R. Schunk, R. Hanbaba and T. Mori, "Interhemispheric comparison of SUNDIAL *F*-region data with global scale ionospheric models," Annales Geophys. 6, 31, (1988).

[7] M. A. Abdu, B. M. Reddy, G. O. Walker, R. Hanbaba, R. W. Schunk and E. P. Szuszczewicz, "Processes in the quiet and disturbed equatorial-low latitude ionosphere: SUNDIAL campaign 1984," Annales Geophys. 6, 69, (1988).

[8] R. G. Roble, R. E. Dickinson and E. C. Ridley, "Global circulation and temperature structure of thermosphere with high-latitude plasma convection,"J. Geophys. Res. 87, 1599 (1982).

[9] R. G. Roble, and E. C. Ridley, "An auroral model for the NCAR thermospheric general circulation model (TGCM)," Annales Geophysicae 5, 369 (1987).

[10] J. C. Foster, J. M. Holt, R. G. Klusgrove, and D. S. Evans, "Ionospheric convection associated with discrete levels of particle precipitation," Geophys. Res. Lett. 13, 656 (1986).

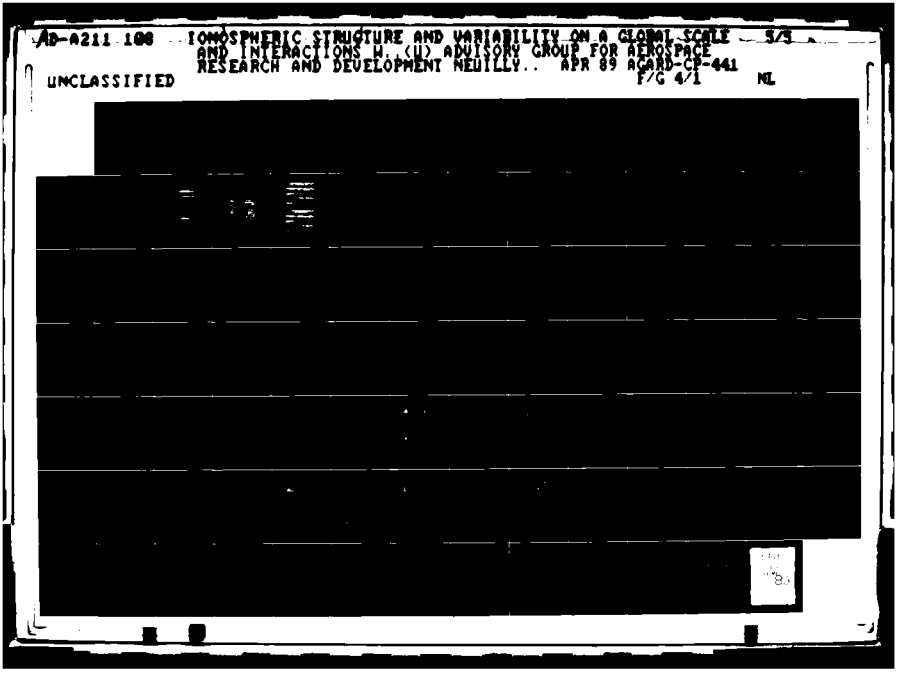
[11] B. G. Fejer, "Equatorial ionospheric electric fields associated with magnetospheric disturbances," in Solar-Wind-Magnetosphere Coupling, Y. Kamide and J.A. Slavin, eds. (Terra Scientific Publishing Co., Tokyo, 1986), p. 519.

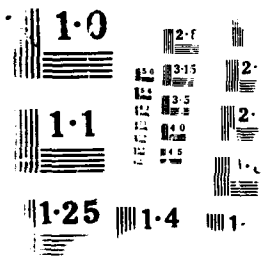
[12] B. G. Fejer, "The equatorial ionospheric electric fields," J. Atmosp. Terr. Phys. 43, 377 (1981).

[13] M. C. Kelley, B. G. Fejer and C. A. Gonzales, "An explanation for anomalous equatorial ionospheric electric fields associated with a northward turning of the interplanetary magnetic field," Geophys. Res. Lett. 6, 301 (1979).

[14] R. A. Wolf, R. W. Spiro and G. A. Mantjoukis, " Theoretical comments on the nature of the plasmapause," Adv. Space Res. 6, 177 (1986).

AD-A211 100 IONOSPHERIC STRUCTURE AND VARIABILITY ON A GLOBAL SCALE — 5/3
AND INTERACTIONS W..(U) ADVISORY GROUP FOR AEROSPACE
RESEARCH AND DEVELOPMENT NEUILLY.. APR 89 AGARD-CP-441
UNCLASSIFIED F/G 4/1 NL




1·0
1·1
1·25
1·4

## DISCUSSION

**B.Reinisch**

How can you dare to derive the instantaneous polar convection pattern from a few isolated observation radars, of which at best two can look into the polar cap?

**Author's Reply**

We expect that our five-station radar study of the high-latitude ionosphere and associated convection electric field should be one of the best studies conducted to date. We are all keenly aware of the importance of the field strength and pattern, the controversy in the dynamics of its morphology, and the synthesis of time-variant satellite measurement of **E** along tracks that has contributed to our current phenomenological prescription of two-cell and multi-cell models. We expect that our study will be part of the overall advance our community is making in understanding the high-latitude convection electric field.

**K.Rawer**

1. The IRI at present uses the official CCIR-map. There was a URSI group discussing recent proposals for improved maps. If CCIR is to follow the proposals of this group, one will introduce a recent Australian program (*Fox et al.*) which has incorporated features introduced by an American group (*Rusch, Anderson et al.*). Unfortunately, the improvement so reached over the oceans is apparently accompanied by not as good agreement as before over the continents.

2. The numerical maps are for monthly averages. Since the full dispersion range is about ±20%, checking with a few days' results is not really conclusive; it is however, quite helpful to know the regions where larger differences have been observed.

3. I agree with you that the present maps are not helpful in the auroral zones and polar caps. Since the physics there differs so much from that at lower latitudes, I doubt whether the present mathematical scheme is adequate; we cannot greatly increase the number of coefficients (which even now is about $10^3$ per map). So my proposal is to establish *separate maps* for the zones and establish a transition zone at the equatorward limit of the oval, or at the trough.

**Author's Reply**

I am glad that this is a comment and not a question.

**G.Wrenn**

You described comparisons of data with model values for an equinox campaign. My experience with single station empirical models is that it is more difficult to make good forecasts of $f_oF_2$, at mid latitudes, near equinox than at solstice. The 'change-over' from summer to winter (or vice versa) patterns occurs near equinox, but the actual timing is uncertain; it seems to depend upon geomagnetic activity. For this reason the SUNDIAL objectives might be better achieved, in the first instance, with solstice campaigns.

**Author's Reply**

Our campaign schedule should nicely evaluate your concerns. Since the October equinox campaign of 1987, we have already conducted campaigns in September/October 1986, June 1987 and March 1988; and we are coming up with a campaign in December 1988, will follow-on every nine months through March 1991. See the February 1988 issue of Annales Geophysicae for a more complete description of measurement rationale and measurement scheduling.

**T.Damboldt**

I would like to comment that low percentage deviations between measured and predicted values of $f_oF_2$ (e.g., in the European area) do not necessarily prove a good agreement. A virtually good agreement is - at least to some degree - favored by an already low numerical value of $f_oF_2$ (in this case 3 MHz). A high percentage difference between predicted and observed values is certainly much more likely if the actual frequencies are high. Perhaps the (percentage) difference together with the standard deviation gives a *more realistic* picture of the agreement between measurements and predictions.

**Author's Reply**

That is a very good observation and I thank you for your comment. As we develop our global-scale comparisons and work toward an adaptive empirical model, we seek to quantify as accurately as possible the degree of agreement/disagreement between the model specifications and the observations. In SUNDIAL science team meetings, we discussed this at some length and advanced the method used in this presentation. The idea in our current format is to superimpose percent differences (positive and negative) directly over global contours of model values for $f_oF_2$. In this way we also can extract the absolute values of differences. This is perhaps the best approach if one is interested in the physics of ionospheric density distributions. On the other hand, your suggestion is likely to be a better choice for propagation channel analysis.

# NORTH-SOUTH TROPICAL F2 LAYER ASYMMETRY AT WEST AFRICAN AND PACIFIC LONGITUDES

by

Colette Davy, Roland Fleury, Pierre Gourvez and Rudi Hanbaba
MER CNET LANNION, B.P. 40
22501 Lannion Cedex
France

and

Paul Vila
CRPE PAB CNET
92131 Issy-les-Moulineaux
France

## ABSTRACT

Because in September-October 1986 the $a_p$ magnetic activity nearly changed with 24-hour steps, this period presents rather simple F2 layer day-to-day changes in the intertropical zone. Along the ten days selected for the equinox Sundial study, the CNET stations: Dakar (17.7°N, 17.2°W, northern subtropic, Ouagadougou (12.4°N, 1.6°W, magnetic equator), Tahiti (17.4°S, 149.2°E, southern tropic) provide a sufficient latitude diversity to characterize the magnetic season asymmetry of generally dominant anti-solar tropical peak densities. On the basis of previous ground-based and satellite meridian-aligned surveys we investigate the F2 layer peak true height and foF2 evolution at these three sites. Low and stationary activity regimes bring about the strongest crest values on the anti-solar side, here estimated from the Dakar variations of September 27, September 28 and October 4 1986. The evolution on our (moderately) disturbed day when the average kp values jumped from 1+ the day before to 4+ show everywhere the same single weaker maximum around 15 LT. Our 1966 complete data along the Tamanrasset median help us investigate the main processes of these two different types of evolution: the intertropical fountain and neutral drag mechanisms are investigated: the respective effects of the $\mathbf{E} \times \mathbf{B}$ dynamo, neutral wind from sub-solar atmospheric heating and from sub-auroral disturbance dynamo are deduced. Coupings between these processes and the conjugate photo-electron forced diffusion are discussed.

## RESUME

Les campagnes Sundial ont relancé l'étude globale de l'interaction entre les systèmes (i) de la dynamo magnétosphérique, couplee par le courant annulaire et les courants sub-orauraux aux deux vortex (ii) des dynamos semi-globales de marée atmosphérique qui se recouplent a l'équateur dans (iii) la dynamo de de l'électrojet équatorial: son champ de polarisation ouest-est alimente le moteur à plasma $\mathbf{E} \times \mathbf{B}$. B couple à la diffusion ambipolaire dans la "fontaine d'Appleton". En période de calme magnetique ce système redistribue le plasma equatorial jusqu'aux deux crêtes tropicales surdenses de la couche F2.

A l'occasion des campagnes de Septembre-Octobre 1986, on retrouve le comportement ci-dessus (malgré la faible densite du reseau d'ionosondes au sol) et les variations d'un jour à l'autre. Ceci est facilité par les variations de l'activite magnetique en periodes homogenes de 24 heures TU.

On evalue l'effet au "tropique reel" de la fontaine equatoriale grâce aux variations typiquement subtropicales de la station de Dakar

On reconstitue la dissymétrie de periode calme due à l'entrainement par le vent neutre grâce au modèles empiriques. On amorce une etude plus fine des couplages internes et mutuel entre les processus classiques (ii) et (iii). Celle du couplage des divers effets de vent neutre avec la dynamo équatoriale et avec la diffusion forcée due aux flux de photo-electrons est proposee, à l'occasion de l'année international de l'electrojet équatorial (I.E.E.Y) 1990—1991.

## INTRODUCTION

The loss twenty years ago of the Tamanrasset tropical ionosonde site should have made the present study impossible; however the magnetic equator has since migrated at a rapid rate on the Atlantic coastal sector, so that the CNET stations Ouagadougou and Dakar have become respectively good equatorial and subtropical sites. We shall make use of the Tahiti data to define the Universal time change in the Southern tropic, and select typical diurnal periods with the reference of previous studies. The main equatorial process, of the fountain diffusion, is generally modified by a semi-stationary neutral wind flow from the subsolar hemisphere side during the diurnal build-up of the intertropical ionization crest system. The day-to-day changes in the equatorial electric field are sampled by the diurnal variations of the two African stations, and the continuity of the wind effect by the comparison with the Tahiti data. For the afternoon hours, we evaluate the counter-effect due to forced thermal diffusion by the unstable situation of asymmetric conjugate photo-electron fluxes on the topside of the winterside crest. This situation contriubutes to an apparent "symmetry" in the statistical results.

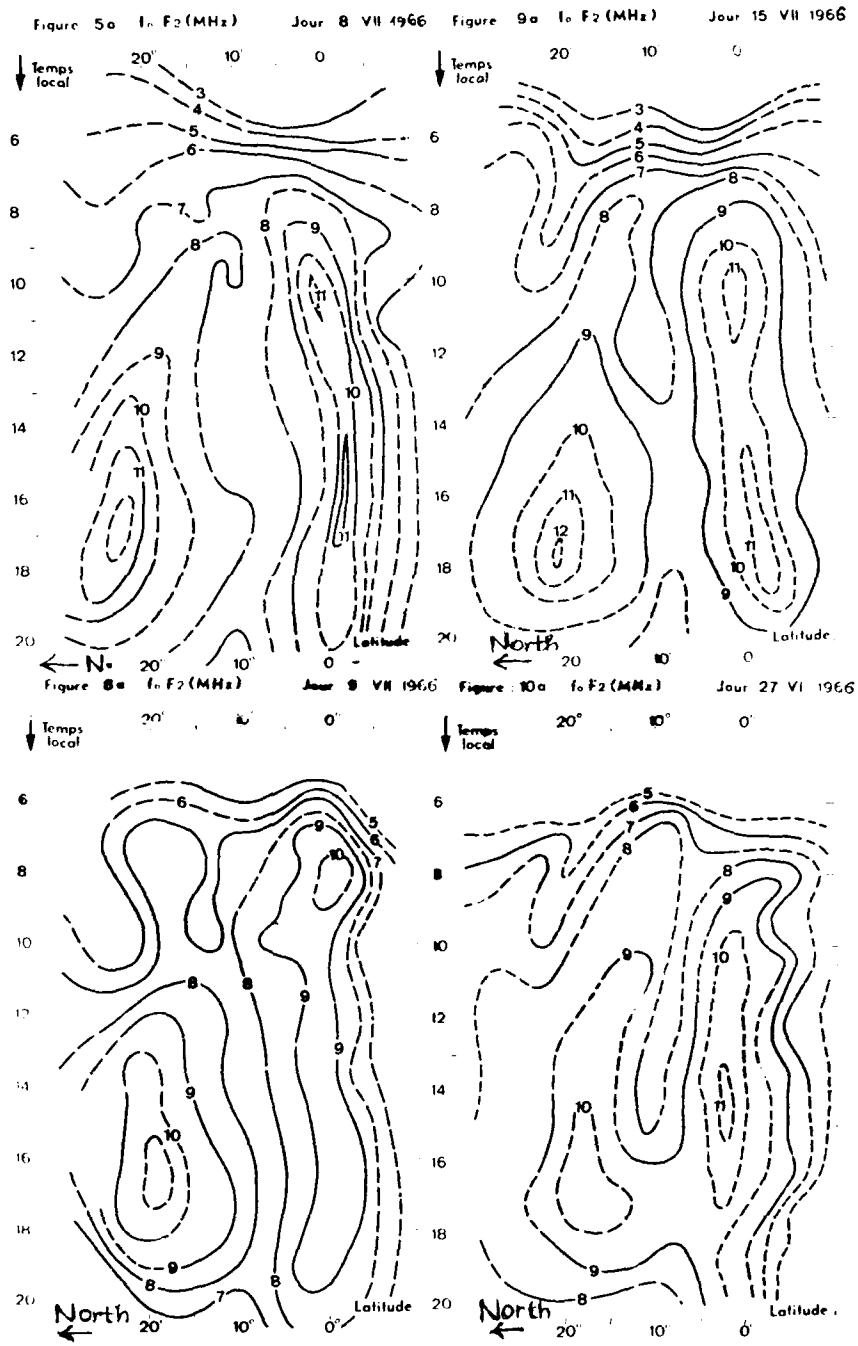

Figure 5a fo F2 (MHz) Jour 8 VII 1966
Figure 9a fo F2 (MHz) Jour 15 VII 1966
Figure 8a fo F2 (MHz) Jour 9 VII 1966
Figure 10a fo F2 (MHz) Jour 27 VI 1966
Temps local
Latitude
North

Figure 1

**The "true" tropical crest**

Figure 1 shows results of a complete survey in June—July 1966 along the Tamanrasset-Kinshasa meridian, with one airborne and eight ground-based ionosondes, which defined the tropical crest with less than one degree of latitude and 15 minute time accuracy. Basically the September solstice situation differs from a mirror image of the June one only by different tidal and neutral wind effects. In the growing phase from 07. to 16. L.T. period, the crest pattern develops more rapidly on the North side in September, and less rapidly in June. Figure 2 shows how to extrapolate the "maximal" density at the effective tropical crest, and its approximate time from the Dakar curves. This will help us characterizing the day-to-day changes of the fountain process.

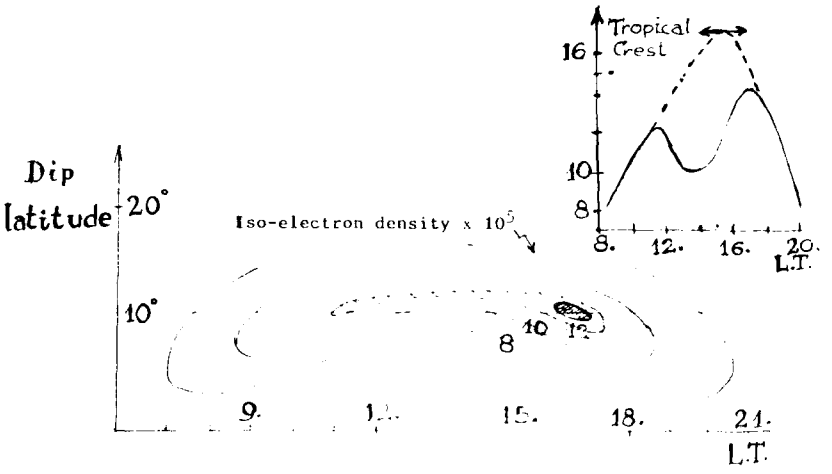



Figure 2

**The fountain changes**

Our typical days are the rather quiet 28 September ($a_m = 22$, Dst = −10) and 4 October ($a_m = 12$, $D_{st} = -10$). The most magnetically active days are 23 and 24 September. On the 23, $a_m = 58$ and Dst = −50 nT. For these two types of days the values of the upward drift $v_z = \mathbf{E} \times \mathbf{B}$ B are almost similar; the ratio of $v_z$ deduced from the H magnetic component at M'bour (dip latitude 9° N) is only 7/8; such a ratio cannot cause the large day-to-day changes observed.

On Table I our estimated maximum tropical crest reaches foF2 = 20 MHz at 14.LT near 15° dip latitude on the two quiet days, as against the single maximum observed on 23 September just over Dakar at 18.LT. At Ouagadougou the corresponding equatorial values are little influenced by $a_p$ activity, with foF2 = 10 MHz around 14.LT for the quiet days and just slightly more at 13.LT on 23 September; on this disturbed day another process must be *contracting* the double-crest structure within the subtropical zone.

**TABLE I**

| Month | Day | $a_m$ | $a_p$ | Dst | DAKAR MHz Peak | DAKAR MHz UT. | DAKAR TROP Crest | OUAGADOUGOU MHz Max | OUAGADOUGOU MHz UT. | OUAGADOUGOU Trough | TAMAN |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 27 | 6 | q0 | 10 | | | | | | | |
| | 23 | 58 | D9 | 50 | 14<br>18 | | 14<br>18 | 10<br>16 | 08<br>13 | 11<br>16 | |
| | 24 | 39 | D7 | 42 | | | | | | | |
| | 25 | 28 | 03 | 10 | | | | | | | |
| | 26 | 40 | E | 25 | | | | | | | |
| | 27 | 33 | q2 | 10 | | | | | | | |
| | 28 | 22 | q2 | 10 | 10<br>10 | 13<br>** | 20<br>14 | 9<br>** | | | 12.4<br>-- |
| | | 16 | q2 | 10 | 9<br>09 | 12<br>16 | 20<br>** | | | | |
| OCTOBER 1966 | 3 | 16 | q2 | 10 | **<br>10 | 12<br>13 | 18<br>15 | 9.6<br>10 | **<br>16 | 08<br>12 | **<br>12 |
| | 4 | 12 | q0 | 10 | **<br>9 | **<br>14 | 23<br>19 | 9.5<br>9 | 18<br>15 | **<br>18 | **<br>13 |

These $N_MF2$ variations from one day to another are abnormally large for the small range of magnetic activity levels. At the effective tropical crest they would be even larger than given by our too-linear extrapolation. This can be explained considering the appropriate phasing of the magnetic activity changes; these mostly arose during the end of UT night intervals. The African sector was therefore favoured for the morning buildup of the fountain mechanism.

The mean $N_MF2$ value is 17.6 MHz with a standard deviation of 3.2 MHz while the Southern crest maximum values averaged on 10.8 MHz with 1.2 MHz standard deviation. These values illustrate the intertropical asymmetry of the African September solstice. Table I gives the complete numerical data.

The double-hump shape of the diurnal variation at Ouagadougou is mainly due to the wide extension of the magnetic tubes of force filled-up by the equatorial Hall drift: we extrapolate the effective equatorial trough under the minimum at Ouagadougou when the bite-out appears too progressive and narrow.

### Neutral wind effects

The electron density asymmetry is more readily estimated with the incomplete African network than real tropical crest maximum amplitude. The asymmetry reaches its maximum on the quiet and moderately disturbed days 24–28 September and 2–4 October, with average Northern crest peak values of 19 MHz in North West Africa, and only 9 MHz on the Southern crest at Tahiti, a density ratio of four to one in favour of the non-subsolar crest; this reveals an intense flow during the whole midday period, from south to north across the magnetic equator.

The asymmetry falls off on the 23 September, our disturbance day; also the equatorial density is high. In this case the cause of the modified more symmetric distribution must be a *convergence* of the neutral wind from both hemispheres near the magnetic equator. $h_mF2$ peak altitudes are seen to rise much earlier in the day and reach somewhat higher; a strong North-South neutral wind flow must be counteracting the northward one from the Southern hemisphere. The disturbance can be attributed to sub-auroral heating in the northern hemisphere; considering the time elapsed between the $a_m$ and $\Delta T$ indices of the auroral disturbance, the average meridian wind flow velocity over the 5000 km from its source to the equator in about 10 hours amounts to 500 km $h^{-1}$, or 140 m $s^{-1}$. This value is consistent with convergence at the magnetic equator which requires gradual mid-latitude damping, to end at the subtropic with about 70 m $s^{-1}$; the mid-latitude storm-heating transport had been suggested from our equatorial results in 1971 (Vila [1]), and theoretically established from incoherent scatter measurements by Richmond in 1979 [2].

### Coupling with conjugate photo-electron thermal diffusion

We have explained [1] on our 1966 data a paradoxal coupling of the neutral wind with fountain development of typical "steady" types of diurnal variation. On the 15 July 1966, illustrated by the upper right panel of our Figure 1 near 15 30 LT, the summerside northern crest near 18° latitude (beware the mis-directed latitude arrow at the bottom right angle of the panels) suddenly develops in density and latitude. The corresponding height pattern showed a *descent* of the crest peak; this contradicts the usual neutral wind drag effect which goes on enhancing (and contracting in latitude) the opposite tropical crest; we showed that this last displacement had sent the opposite crest away from the paths of photo-electrons hitting the topside border of the northern conjugate crest; this deposits a strong heat gradient on the topside border, pushing the crest peak to North down the magnetic lines of force. The same variations take place on 4 October 1986 between 17 and 18 UT (16 and 17 LT), as we see on our Figure 3. The three-peak $f_0F2$ variation at Dakar had been a riddle for the observers since the ionosonde setup in the late forties.

### Thermosphere Coupling Studies

These confirmations of the dynamics get us back to the problems facing modelists and previsionists now. The electrodynamics of the intertropical thermosphere are still empirical, as we have not yet been able to model the effective variations of the electrojet itself from one day to another. The relative couplings of the tidal modes with the interhemispheric neutral wind at solstices needs elucidating by *independent* measurements of the wind velocity on both sub-tropical sides. We also need more detailed evaluations of the electric field and current effects on the inner and outer border of the equatorial electrojet *wings* to understand the couplings of this current with the midlatitude dynamo vortices. The I. T. C. S. group and the I. E. E. Y. campaigns expected by IAGA for 1989 are a great issue, since modelling the ionization for prediction purposes would be the first positive use of our science to the people living in the tropical Sahel.

## CONCLUSION

The dynamic processes of the intertropical F2 layer demonstrated from results of 1965 and 1966 complete surveys [2], are confirmed by our September 1986 Sundial series, highly typical of moderate Sunspot activity conditions.

1 – The very weak $\mathbf{E} \times \mathbf{B}$ equatorial fountain changes are found unable to explain the intense day-to-day modulation in the ionic density of the antisolar northern crest.

2 – The North-South asymmetry of the tropical crests deduced from the Dakar and Tahiti variations *on quiet days* reveal the major influence of the trans-equatorial neutral wind ion drag.

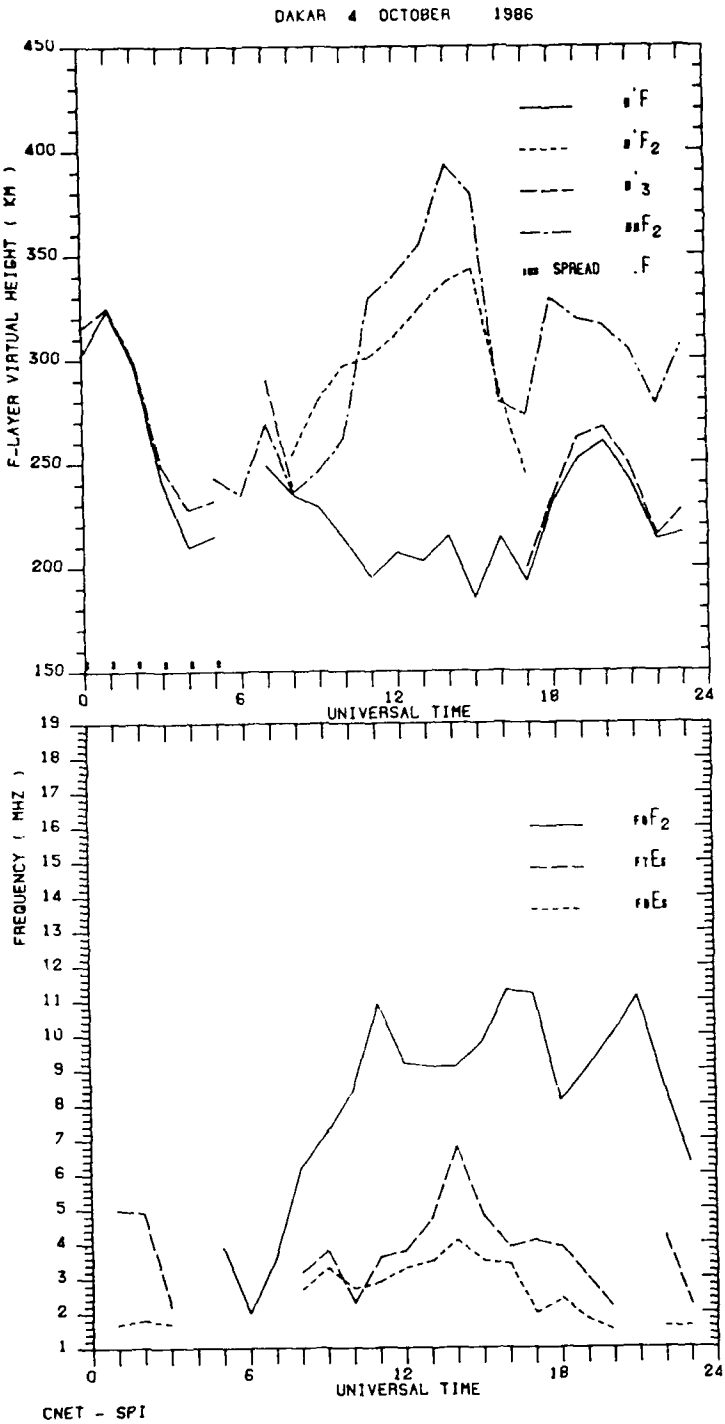

DAKAR 4 OCTOBER 1986
F-LAYER VIRTUAL HEIGHT ( KM )
450
400
350
300
250
200
150
0
6
12
18
24
UNIVERSAL TIME
SPREAD .F
FREQUENCY ( MHZ )
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
UNIVERSAL TIME
CNET - SPI

Figure 3

3 — On weakly disturbed day 23 September 1986 the reduced density and latitude excursion of the northern crest illustrate the effect on that day in the African sector of *neutral wind flows converging* at the magnetic equator; the North south flow opposed to the uniform transequatorial wind is due to auroral magnetic substorm activity (Gourvez [1] and Richmond [3]).

4 — Another typical process of *forced-diffusion* can affect the afternoon antisolar crest during quiet days: this crest rapidly enhances in ion density, subsides in height and travels further away from the dip equator without changing magnetic tube of force; the neutral wind asymmetry, leaving the subsolar crest *inside* the subtropical zone enables *intense photo-electron fluxes* to reach the conjugate antisolar crest on its topside border: there they develop a negative electron temperature gradient which forces the whole crest further, until the electron density gradient is sufficiently lowered by the diffusive transport. In low and moderate Sunspot activity periods, such events last generally less than one hour. They produce at subtropical latitude a sharp *third peak* in the FoF2 variation.

5 — Future improved simulations could use data planned for the *1990—1991 I.E.Y.* Campaign; this will bring finer-scale *simultaneous* measurements of the parameters modulating neutral wind electrodynamic changes near the magnetic equator.

**Acknowledgements**

We thank the workers of the MER-CNET-Lannion groups who, under Dr Jacques Papet-Lepine maintained the low-latitude CNET ionosonde network for the last 20 years. The true height conversion of ionogram profiles by the Titheridge method was programmed conveniently from its early stages at Lannion by Dr Patrick Lassudrie Duchene.

## REFERENCES

[1] Gourvez P. Thèse 3ème Cycle, Etude de la couche F equatoriale à partir des données des satellites ISIS et de stations magnétiques au sol, Paris VI, 1972.

[2] Vila P.M., New dynamic aspects of intertropical F2 ionization, Radio Sci. 6 (11) 945—956.

[3] Richmond A.D., Thermospheric heating in a magnetic storm: dynamic transport of energy from high to low latitudes, J.Geoph. Res., 85, 5259, 1979.

## DISCUSSION

**K.Bibl**

I see no justification to extend the noontime $f_0F_2$ for Dakar to 20 MHz. All the data shown from the 1966 campaign have a rather flat maximum during daytime on all stations, at the magnetic equator as well as at the ionization crest.

**Author's Reply**

There are days when the change in the electrojet field and in the fountain regimes would not allow this. I am relying on the fact that this period was rather a period of steps in the magnetic activity. I think these high values are comparable to what we derived in 1966.

# JOINT VHF COHERENT RADAR AND OBLIQUE SOUNDER OBSERVATIONS FOR SHORT TERM IONOSPHERIC FORECASTING

by

P. S. Cannon*
M. Lester+
E. C. Thomas+
A. H. Dickson*
T. B. Jones+
R. Maude+
P. J. Rylah+
A. K. Shukla*

*Radio and Navigation Department
Royal Aircraft Establishment
Farnborough
Hampshire GU14 6TD UK

+Ionospheric Physics Group
Department of Physics
University of Leicester
Leicester LE1 7RH UK

SUMMARY

An experiment which utilises two VHF coherent radars located in northern Scotland together with an oblique chirp sounder, operating over a path from Norway to northern Scotland is described. The expected sounder path falls within the radars field-of-view. The paper describes some initial results from the first in a series of experiments. A statistical correlation is found between the occurrence of radar backscatter and the occurrence of sporadic-E. On some occasions a positive correlation also exists between the backscatter signal to noise ratio and the highest frequency measured from the sporadic-E trace on the ionogram. On other occasions the correlation is poor. An initial interpretation of these data is presented which suggests that propagation of the HF sporadic-E signals via a simple plane ionosphere is inadequate to explain the observations.

## 1 INTRODUCTION

Obliquely propagating high frequency (HF) radio signals are subject to a variety of propagation effects, which include attenuation, reflection, refraction, diffraction, scatter, polarization changes, multi-path propagation and ducting[1]. At high latitudes some, if not all, of these effects are accentuated by the dynamic nature of the coupled ionosphere-magnetosphere system.

The electron density in the E-region of the ionosphere is controlled by (a) the presence of a time varying convection electric field of magnetospheric origin and (b) the presence of intense particle precipitation in regions surrounding the geomagnetic poles, the auroral ovals[2]. Predicting the behaviour of the high latitude ionosphere on a short time scale (1 day or so) remains an important problem for communications and over the horizon radar applications. Existing empirical models of the high latitude ionosphere[3] are of limited benefit for prediction purposes because of the highly dynamic nature of the coupled ionospheric-magnetospheric system. Consequently, more emphasis is now being placed on developing models which are based on a physical understanding of solar terrestrial relationships[4]. Before such models can be employed directly, however, several outstanding questions concerning HF propagation at high latitudes need to be answered.

In the high latitude ionosphere the incidence of E-region irregularities is high and one outstanding problem is the effect of E-region ionospheric irregularities on, and their relationships to, obliquely propagating HF signals. Elkins[5] describes a predictive model for high frequency radar auroral clutter based upon the occurrence statistics and morphology of sporadic-E[6,7]. Implicit in this technique is the assumption that the same ionospheric irregularities that give rise to vertical incidence high latitude sporadic-E also give rise to oblique incidence HF auroral clutter. Elkins regards this as a reasonable starting point in view of the similar sounding frequency ranges involved in observing both phenomena. Elkins considers that a direct relationship between HF sporadic-E and VHF backscatter is more questionable given that different plasma instabilities might be dominant at the two different wave numbers. However, a VHF radar system has the advantage that it is able to unambiguously determine the location of the E-region irregularities since it is unaffected by refraction. Furthermore theoretical investigations suggest that small scale irregularities are produced by the cascading of large scale irregularities to smaller wavelengths.

To further investigate this relationship between HF sporadic-E and VHF backscatter a series of co-ordinated HF and VHF experiments are being undertaken. The same region of the ionosphere is observed by two VHF coherent auroral radars, which measure the backscatter from E-region irregularities of 1 m scale size, and an oblique HF sounder which

determines the HF propagation parameters. The sounder path includes the auroral and sub-auroral ionosphere and the one and two hop reflection points lie in, or close to, the field-of-view of the radars. This paper presents the initial results of the first in this series of experiments.

## 2 INSTRUMENTATION

### 2.1 The radars

During the experiment, two auroral radars were operated from Wick, Scotland. The Wick station of the Sweden and Britain Radar Experiment (SABRE) consists of eight receiver beams each with angular width of approximately 3.5 degrees, giving a total beam width of approximately 28 degrees (Fig 1). Coherent backscatter is measured in the range interval 495 km to 1230 km along each beam, with a temporal resolution of 20 seconds in Mode 1 and 1 minute in Mode 3. Apart from the temporal resolution, the major difference between the two modes is that Mode 3 measures a more detailed radar spectrum[8], although this information is not employed in this paper.

The other radar which operated during the campaign was the SABRE Altitude Determining Interferometer Experiment (SADIE)[9]. This radar is an extension of SABRE and employs two parallel, horizontal, colinear antennas separated by 200 m. These antennas have an azimuthal beam width of 3.6 degrees and are orientated so as to provide a beam 26.8 degrees east of north thus overlaying beam five of SABRE (Fig 1). By measuring the phase difference of the backscattered radar echoes received at the two antennas, the angle of elevation of the signals may be determined at each range. From this information the mean height at which the radar signals are scattered, within each range cell, may be calculated. The accuracy obtained with this system varies with range but is approximately ±1 km at ranges less than 1000 km which includes the observations reported in this paper. SADIE also operates in the two modes described for SABRE and both radar systems employ 100 μs pulse widths providing a range resolution of 15 km.

Both radars are sensitive to 1 m scale auroral irregularities, which are generated by two different mechanisms, the gradient drift instability[10] and the two stream instability[11]. Each process dominates in a different regime of the relative electron-ion velocity, with the two stream instability dominating when this velocity is larger than the local ion-acoustic velocity[12]. In both cases, linear theory predicts that the unstable irregularities will propagate with a velocity equal to that of the electrons. The Wick radars measure the line of sight velocity in each beam as a function of range and time.

The received backscatter power is also dependent upon the strength of the ionospheric current[13], although there is a threshold current which must be exceeded before the instabilities can be generated. The backscatter power, together with the height of the irregularities are the two main radar parameters in this study.

### 2.2 The oblique sounder

An oblique, frequency modulated, continuous wave (FMCW) ionospheric sounder was operated for about 12 hours every day during the experiment. The 100 watt sounder transmitter was located at Bodø in northern Norway (67.27N, 14.39E, geographic). The transmitter antenna was a horizontal log-periodic covering the band 4 to 32 MHz with an azimuthal beam width of approximately ±40 degrees centred within 4 degrees of the great circle bearing to Wick. The antenna provided about 8 dBi of gain over the band at elevation angles between 35 and 40 degrees. At lower elevation angles pertinent to two and one hop E-layer reflections (approximately 17 and 5 degrees respectively) the gain reduced to approximately 5 dBi and -5 dBi respectively. These correspond to a 10:1 power advantage for the two hop path relative to the one hop path.

The sounder receiver was installed alongside the coherent radars with an antenna which consisted of a wideband, horizontally polarised sloping vee antenna beamed towards Bodø. This antenna was frequency dependent and again the antenna favoured two hop modes. The sounder's characteristics are summarised in Table 1. The sounder only measures relative group delay as no absolute timing is maintained by either the transmitter or receiver. The receiver was interfaced to a Hewlett Packard Series 200 computer for experiment control and ionogram storage. The conventional oblique ionograms (Figs 2 and 3) are supplemented by various additional information. The top panel displays signal quality, a parameter which is relevant to the communications engineer, the second panel the automatic gain control voltage and the third panel the signal strength in dB, with a maximum value of 60 dB. The panel to the left of the ionogram gives system control and data logging details.

Table 1

Sounder parameters

| Parameter | Value |
|---|---|
| Sweep range | 2-30 MHz |
| Sweep rate | 100 kHz $s^{-1}$ |
| Sweep time | 280 s |
| Baseband bandwidth | 500 Hz |
| Number of baseband spectral estimates | 200 |
| Maximum relative group delay measured | 5 ms |
| Maximum virtual distance measured | 1500 km |
| Relative group delay resolution | 25 μs |
| Virtual range resolution | 7.5 km |

3 RESULTS

The experiment was conducted from 1 December to 11 December 1987 (days 335 to 345 inclusive). Periods when the sounder or radars were non-operational are illustrated in Fig 4 as hashed areas.

Fig 4 also illustrates those intervals when the sounder detected sporadic-E modes and when the radars detected auroral backscatter. On some occasions the received HF sporadic-E signals were low in amplitude and exhibited a top frequency as low as 5 MHz, but on other occasions the signal strength was some 60 dB higher, with a top frequency greater than 30 MHz. In the following text we shall refer to these modes as being due to a sporadic-E layer. This generic title includes, however, a multitude of different signal types ranging from those with small dispersion (Fig 2) to those with large dispersion (Fig 3) and probably includes scattered and refracted modes. These two ionograms are separated by only 15 minutes and clearly indicate the temporal variability of the layer. Fig 4 displays those periods when backscatter occurred anywhere in SABRE beam 5 and SADIE beam 1. These two beams overlap but lie to the west of the great circle path between Bodø and Wick. Data from SABRE beam 7, which includes the path midpoint have also been examined for certain periods but are not presented here because they do not relate to coincident SADIE height data. Limited comparisons of the returns from SABRE beams 5 and 7 show that although there are some detailed differences between the returns in each beam there are few gross variations which would significantly affect Fig 4.

The radar and oblique incidence signals are correlated in a statistical sense but on some occasions radar returns exist without sporadic-E and vice-versa. In order to help understand why this is so and to investigate this further a small period of the data, 14 UT (day 344) to 03 UT (day 345) was selected for detailed analysis. Day 344 was the most geomagnetically active day of the campaign.

Figs 5 and 6 are SABRE and SADIE range-time-intensity (RTI) plots covering these periods; the close correspondence between the two diagrams indicates that both systems are looking at the same region of the ionosphere. Significant returns exist in the periods 14-21 UT, *i.e.* in the afternoon convection cell and again between 0030 and 0230, *i.e.* in the morning convection cell.

Fig 7 illustrates the variation of the top frequency of the sporadic-E trace (FtEs) with time, together with the variation in SADIE signal to noise ratio (SNR), at a range of 930 km from Wick. For propagation on a great circle path 930 km is close to one of the two hop (plane ionosphere) reflection points. It is also close to a one hop reflection point if either a large scale north-south ionospheric tilt of 3.5 degrees, or a smaller scale equivalent electron density latitudinal gradient, is postulated. The correlation of FtEs with the SNR is clear within the period 15 UT to 19 UT, with both the SADIE SNR and FtEs rising three times in synchronisation.

Whilst the temporal variations of FtEs and SNR are similar, the amplitudes are not as closely related. This is particularly evident in the period 17 UT to 19 UT (Fig 8, middle and bottom panels). After 1830 FtEs increases to nearly 30 MHz, compared to 12 MHz at 1715, whereas, the SADIE SNR after 1830 peaks at the same value (35 dB) as at 1715 UT. Furthermore after 1830 there is a significant deviation in the height of the backscatter determined by SADIE (Fig 8, top panel) from the normally accepted scattering altitude of around 110 km[14].

After 19 UT the sporadic-E continues (Fig 7) but the SADIE backscatter amplitude decreases. This reduction is associated with a small bite out in the irregularity intensity (Figs 5 and 6, at 930 km) around 1945 UT. Subsequently, the SADIE SNR increases again after 2015 UT (Fig 7).

The SADIE SNR eventually falls off at 2050 and does not return again until 0115, but throughout this time the sporadic-E continues at a very strong level. After 02 UT no SADIE returns are seen but the sporadic-E continues albeit with decreasing strength and decreasing FtEs.

The mid-point of the sounder path lies at a range of 650 km from Wick and SADIE SNR data at this range has also been examined over the interval 14 UT (day 344) to 03 UT (day 345). Although not illustrated here the correlation of SADIE SNR at 650 km range is less well correlated with FtEs than the observations at 930 km.

## 4 DISCUSSION

An attempt has been made to investigate a correlation between VHF backscatter returns and HF sporadic-E propagation. Coherent backscatter is measured by the radars only when the convection electric field exceeds a certain threshold and when the resulting electrojet current system falls within the radar range gates. The strongest backscatter is statistically associated with the main electrojets[15] and only weak backscatter is found near the convection reversals[16]. Because of the location of the Wick radars the electrojets lie to the north of the radar viewing area during most mornings and no returns are observed. Both a convection electric field and adequate ionospheric conductivity are required to sustain the currents and the latter may be enhanced by particle precipitation. High latitude sporadic-E layers are caused by a variety of phenomena including particle precipitation and wind shears etc. One common factor between VHF radar backscatter and sporadic-E is particle precipitation. From a statistical viewpoint, therefore, a correlation between the two phenomena may be expected. However, at the moment we cannot rule out other processes.

It is important to address the question of why any correlation exists between the SADIE SNR and FtEs. Clearly the correlation is extremely good between 15 UT and 1845 UT. Jones *et al*[16] give the backscatter cross section, $\sigma$, as:

$$\sigma \propto N^2 \langle (\Delta N/N)^2 \rangle \qquad (1)$$

where $N$ is the mean electron density and $\langle (\Delta N/N)^2 \rangle$ is the mean square of the small scale relative electron density fluctuations with wavevector $2k$, where $k$ is the radar vector. Taking $\langle (\Delta N/N) \rangle$ to be a slowly varying function of $N$, a first order approximation[17] gives:

$$\sigma \propto N^2 \quad . \qquad (2)$$

If the HF signals are assumed to be reflected, the top frequency of the sporadic-E trace is controlled by the plasma frequency. Since $N$ is proportional to the square of the plasma frequency it follows from equation (2) that:

$$\sigma \propto (\mathrm{FtEs})^4 \quad . \qquad (3)$$

For a constant background noise level a higher SNR might, therefore, be expected when FtEs increases, which is consistent with the data between 15 UT and 1845 UT. These ideas, however, require further study.

Possible path geometries for the HF sporadic-E modes are now examined. A broad correlation in time, and to a lesser extent, amplitude exists between the SADIE SNR at 930 km range and FtEs for part of the period under consideration. A weaker correlation exists with SADIE SNR data at 630 km range, which is midway to Bodø from Wick, and this suggests that reflection of the HF signals is not over a one hop path. To achieve a two hop path with top frequencies as high as 30 MHz, requires a very high electron density both 400 km and 900 km north of Wick. A 1E+1F mode cannot be always discounted, especially during the early afternoon when F-region modes are observed. It is, however, unlikely that the 1E+1F mode is the dominant mode in this interval, since the F modes are generally very weak and at no time does the 1F2 mode support signals above 10 MHz. The possibility of a 2E mode cannot likewise be ruled out but this would require a single strong patch of sporadic-E, sometimes capable of reflecting signals up to 30 MHz, covering some 500 km of latitude, or two such regions of sporadic-E at the reflection points. The 2E mode is also strongly attenuated by D-region absorption.

A further possibility, that of a tilted ionosphere, cannot be as easily dismissed. A likely position for this reflection point is suggested by the temporal correlation between FtEs and the SADIE SNR to be close to a point 930 km from Wick. This mode would require, however, very low angle elevation coverage by the Wick HF sounder receiver antenna. Assuming an altitude of 110 km for the sporadic-E region the take-off angle at Wick would be only 2.6 degrees. At this elevation angle the receiver antenna is very inefficient.

The discussion has so far centred on reflection of the HF signals, however, scattering of the radio waves may also occur for all or some of the time[18]. If this is so FtEs will be dependent on the receiver SNR. One way that the HF SNR could increase is for the altitude, from which the HF signals are scattered, to increase since this will be accompanied by an increase in receiver antenna efficiency. Such a height increase in the radar backscatter does take place at 1830 UT and it is accompanied by an increase in FtEs.

The HF modes, like the VHF modes, may specifically have been scattered from field aligned irregularities. The field line orthogonality contours for a transmitter at Bodø and receiver at Wick are illustrated in Fig 9 for a height of 105 km. The paths which most nearly satisfy the orthogonality condition lie in a region about 1000 km from Wick where the best correlation between the SADIE SNR and FtEs is observed.

Between 21 UT and 01 UT the radars most likely observe the reversal in the convection pattern and as a consequence the low backscatter amplitude is to be expected. At certain other times the lack of correlation might be explained by the movement of the electrojet to latitudes polewards of those viewed by SABRE and SADIE. This would be particularly likely in the post midnight period. Furthermore, the lack of correlation may be due to the sporadic-E being caused by a mechanism other than particle precipitation.

## CONCLUSIONS

This paper describes some initial results from joint coherent radar and oblique sounder experiments.

A good correlation between the VHF backscatter SNR and the HF FtEs is achieved over long periods but there are other times when the correlation degrades. Various reasons for this have been suggested. The likely position for the HF reflection point has been determined from a comparison with the SADIE backscatter data and this is about 930 km from Wick. Preliminary analysis indicates that one or two hop HF modes from a simple plane ionosphere are unlikely. Two other explanations are discussed involving a tilted ionosphere and scattering rather than reflection.

This investigation has attempted to establish the usefulness of VHF radars as remote sensors of sporadic-E and hence to apply them to ionospheric forecasting. The results are so far inconclusive. In the future a more detailed analysis (e.g. correlation studies along the radar beams at various ranges) and more sophisticated experiments (e.g. employing vertical incidence ionosondes for diagnostic purposes) are required. Even though the desired objective of applying coherent radars as sensors of sporadic-E may not be achieved it is clear that this series of experiments will produce new and interesting information on the morphology of high latitude sporadic-E and its relationship to auroral irregularities.

## REFERENCES


1 Thrane, E.V. Media effects on systems in the high latitude region, Paper 1.2, AGARD Conf Proc 382, Propagation Effects on Military Systems in the High Latitude Region, Alaska, June 1985

2 Quegan, S., Milsom, J.D. and Herring, R.N, Ionospheric modelling and HF radio systems at high latitudes, Paper 3.1, AGARD Conf Proc 382, Propagation Effects on Military Systems in the High Latitude Region, Alaska, June 1985

3 Elkins, T. and Rush, C. A statistical predictive model of the polar ionosphere, In 'An Empirical Model of the Polar Ionosphere', Editor T. Elkins, AFCRL-TR-73-0331, Air Force Cambridge Research Labs, Bedford, Mass, USA, 1973

4 Quegan, S., Gill, R.S. and Lockwood, M, Theoretical modelling of high latitude ionospheric densities observed by EISCAT, *J. Atmos. Terr. Phys.*, In Press 1988

5 Elkins, T.J, A model for high frequency radar auroral clutter, Rome Air Development Center, Report RADC-TR-80-122, 1980

6 Besprozvannaya, A.S. and Shchuka, T.I, Space-time distribution of Es formations associated with visible auroral forms, *Geomag. and Aeron.*, 6, 878, 1971

7 Besprozvannaya, A.S. and Shchuka, T.I, Sporadic ionization of the ionospheric E-region at high latitudes as a function of magnetic activity, *Geomag. and Aeron.*, 12, 401, 1972

8 Schlegel, K., Thomas, E.C. and Bahle, P. A statistical study of auroral radar spectra obtained with SABRE, *J. Geophys. Res.*, 91, 13483-13492, 1986

9 Thomas, E.C. SADIE, Leicester University Ionospheric Physics Group Technical Report 42, 1987

10 Rogister, A. and D'Angelo. N, Type II irregularities in the equatorial electrojet, *J. Geophys Res*, 3879-3887, 1970

11 Farley, D.T, A plasma instability resulting in field-aligned irregularities in the ionosphere, *J. Geophys Res*, 65, 6083-6097, 1963

12 Fejer, B.G. and Kelley, M.C, Ionospheric irregularities, *Rev Geophys Space Phys*, 18, 401, 1980

13 Burrage, M.D, Radar studies of high latitude convection flow, PhD Thesis, University of Leicester, 1988

14 Abel, W.J. and Newell, R.E, Measurements of the afternoon Radio Aurora at 1295 MHz, *J. Geophys Res*, 74, 231, 1969

15 Greenwald, R.A., Ecklund, W.L. and Balsey, B.B, Radar observations of auroral electrojet currents, *J. Geophys Res*, 80, 3635, 1975

16 Jones, T.B., Waldock, J.A., Thomas, E.C., Stewart, C.P. and Robinson, T.R, SABRE radar observations in the auroral ionosphere, Paper 6.4, AGARD Conf Proc 382, Propagation Effects on Military Systems in the High Latitude Region, Alaska, June 1985

17 Robinson, T.R, Leicester University, UK, Private Communication, 1988

18 Dudeney, J.R. and Roger, A.S, Spatial structure of high latitude sporadic-E. *J. Atmos Terr Phys*, 47, 6, 529-535, 1985

ACKNOWLEDGEMENTS

We extend thanks to Capt W. Johansen, of the Norwegian Armed Services, for arranging the deployment and running of the sounder transmitter in Norway. ECT and RM acknowledge support by the SERC under Contract SGD/10696 and PJR acknowledges the support of the SERC for a CASE studentship.

*Copyright, ©, HMSO, London 1988*

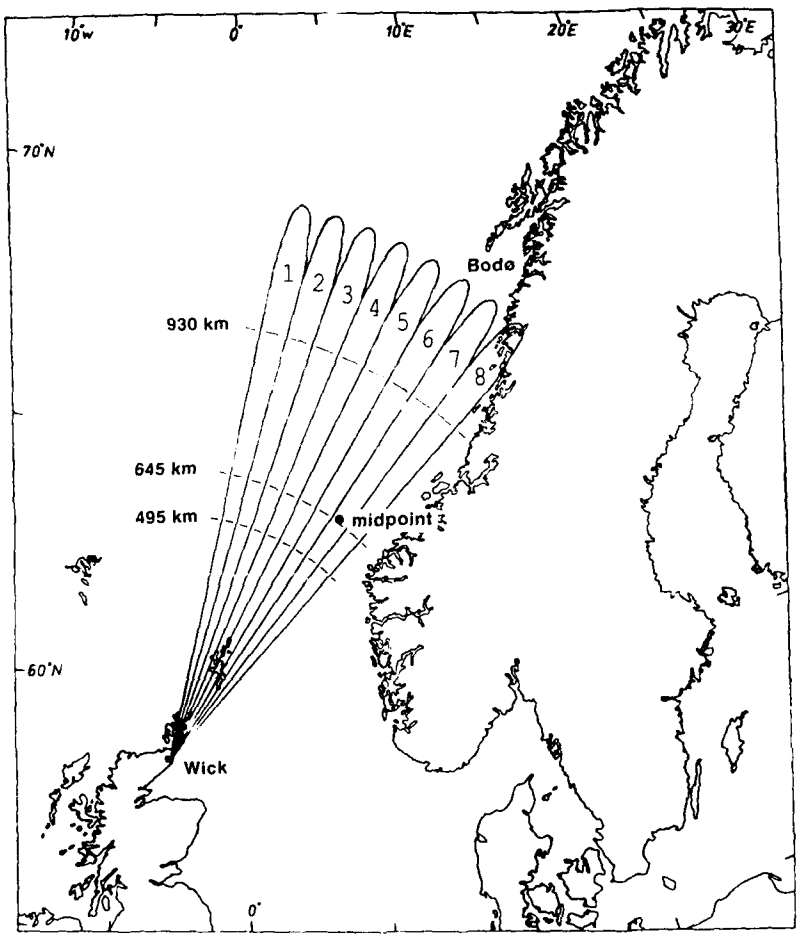



Fig 1 Beam and path geometry of the Wick radars and the oblique sounder

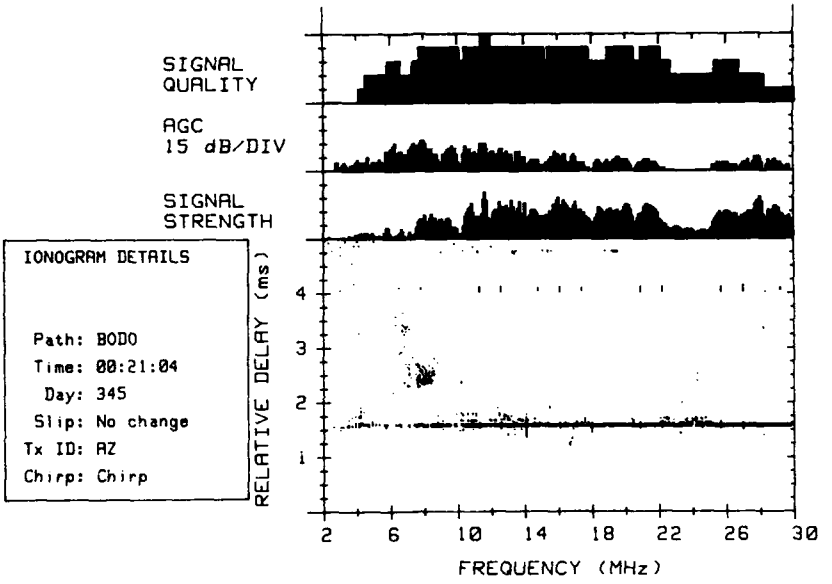



Fig 2 An oblique ionogram for the Bodø Wick path with a sweep start time at 0021 on day 345

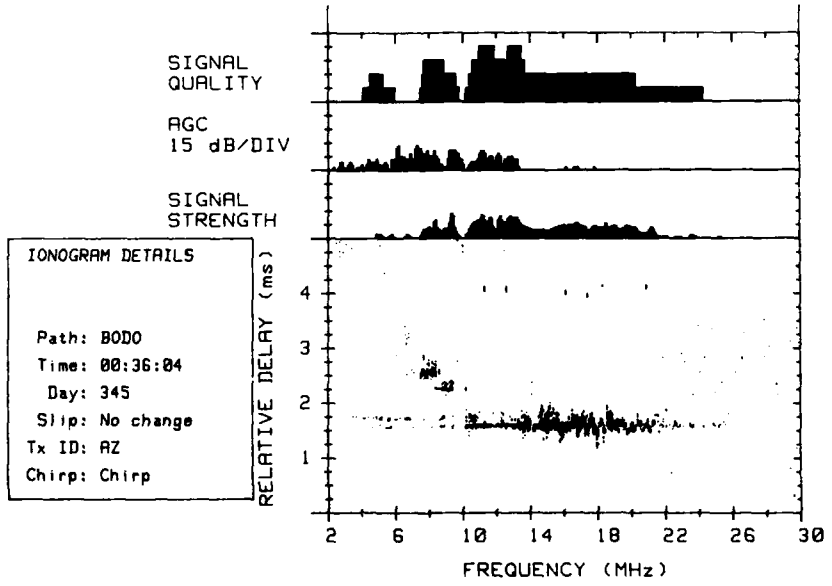



Fig 3 An oblique ionogram for the Bodø Wick path with a sweep start time at 0036 on day 345

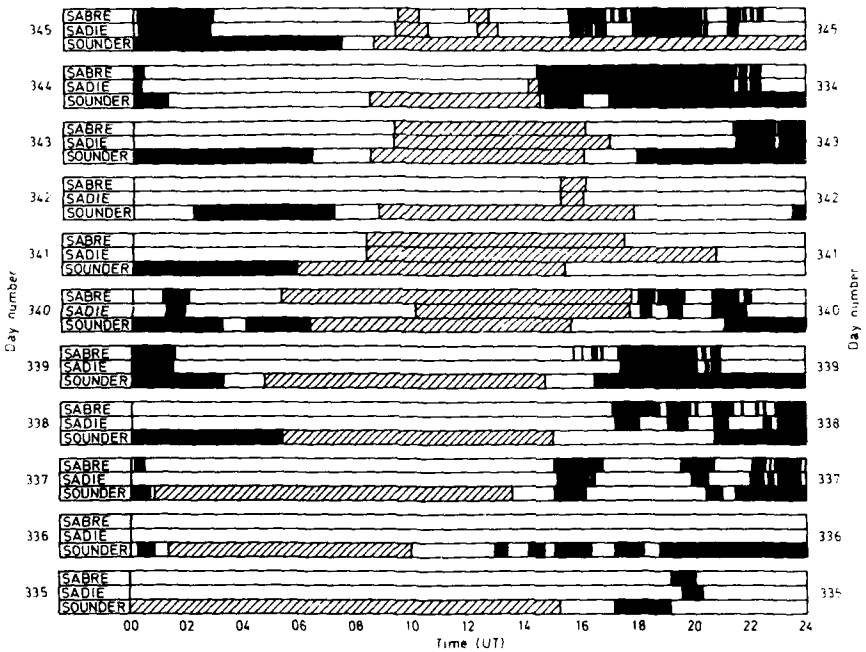



Fig 4 Correlation of SABRE and SADIE radar returns with oblique incidence sporadic-E modes

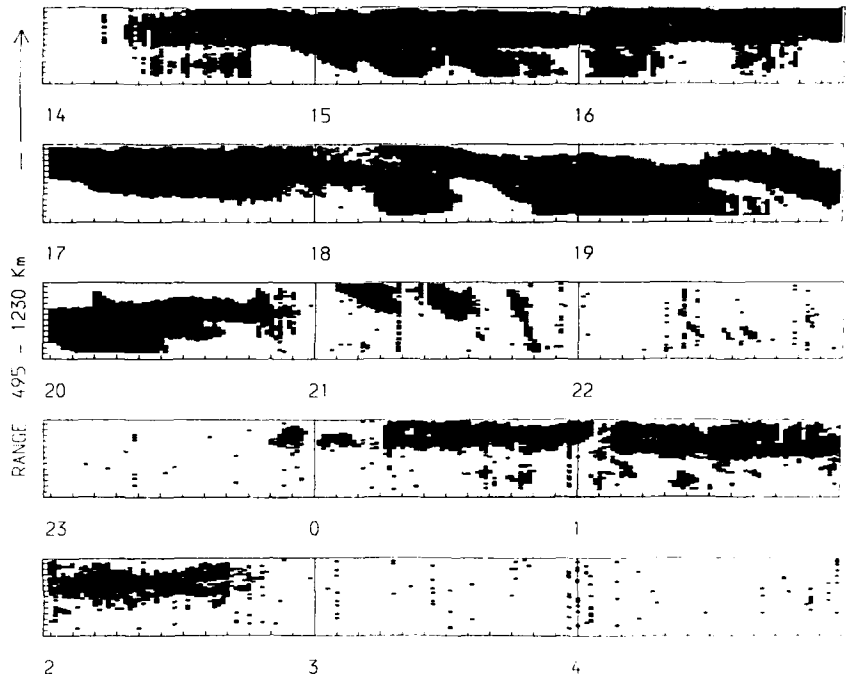



Fig 5 Range-time-intensity plot for the SABRE radar for part of days 344 and 345

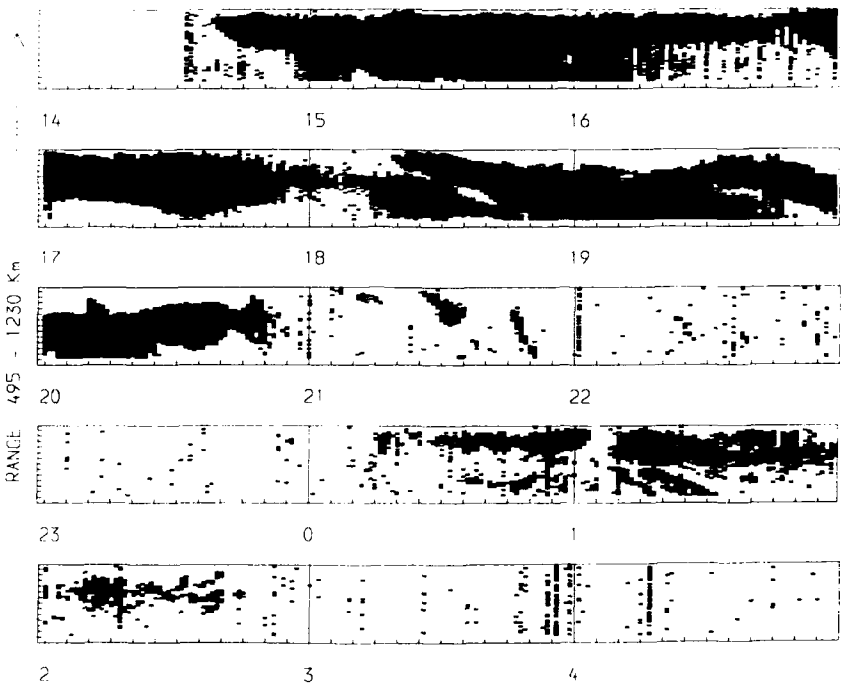



Fig 6 Range-time-intensity plot for the SADIE radar for part of days 344 and 345

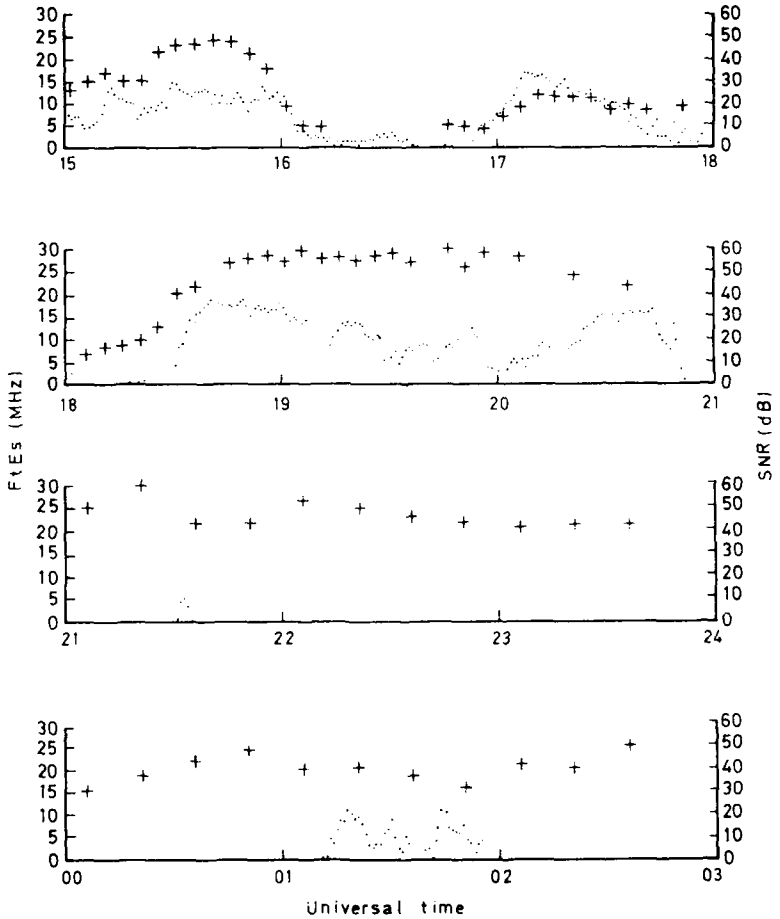



Fig 7 Temporal variations on days 344 and 345 of the highest propagating sporadic-E frequency (FtEs) +++++ and the signal-to-noise ratio (SNR) ..... of SADIE radar returns

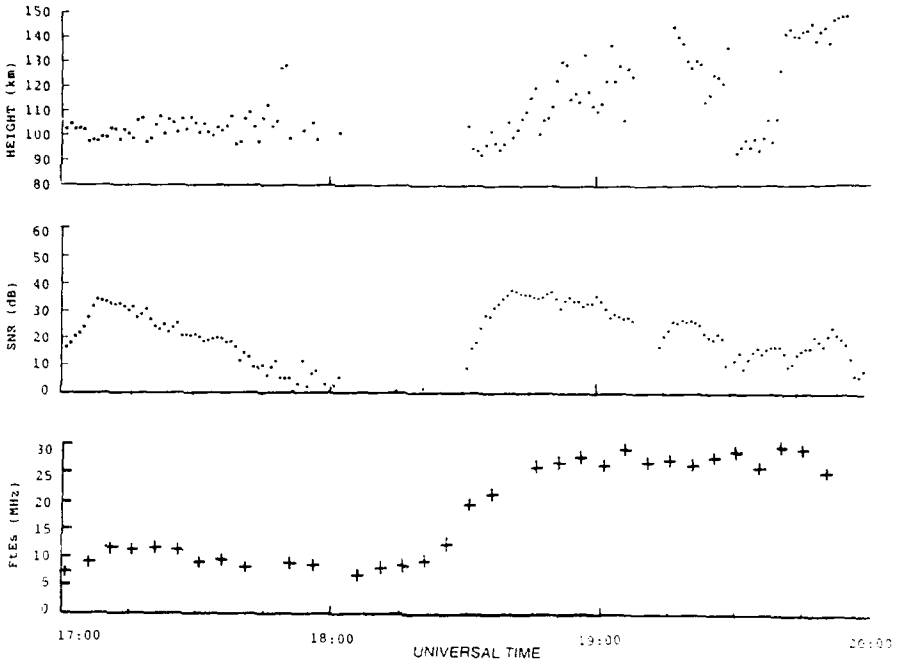



Fig 8 Irregularity height (top), SADIF backscatter power (middle) and FtEs (bottom) for the period 17 UT to 20 UT [illegible]

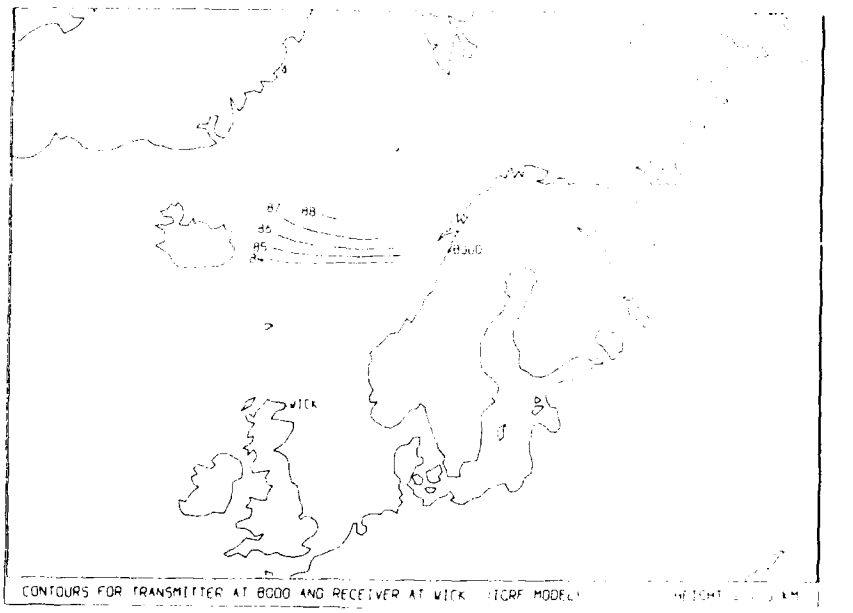



Fig 9 Orthogonality contours (assuming no refraction) for a transmitter at Bodø and a receiver at Wick

## DISCUSSION

**C.Haldoupis**

First of all, I would like to say that you have presented a very interesting set of data. I would, however, like to add you should try to apply the gradient drift instability, as a possible mechanism, in explaining your results. Sporadic $E$ is associated with a very large vertical electron density gradient which you would expect to have a sizeable component perpendicular to **B**. This, in conjunction with a small $E$-field component pointing in the same direction, would destabilize the plasma and cause eventually, through secondary instabilities, detectable short wavelength waves. The fact that you have at times no correlation could be because the gradient drift conditions fail. This proposition can be consistent also with the data shown in Figure 8.

**Author's Reply**

To investigate this suggestion, we require information concerning the vertical electron density distribution, which is not available with the experimental configuration described. In order to find additional information, the experimental geometry for future campaigns is to be changed to provide data from a vertical ionosonde beneath the HF sounder midpath, and from EISCAT. With these additional data, it should be possible to obtain a more detailed interpretation of the connection between HF sporadic-$E$ and VHF radar backscatter.

# RADIO PROPAGATION CONDITIONS FORECAST ABOVE A LOW LATITUDE REGION

H. SIZUN
M. LISSILLOUR

C.N.E.T.
Département LAB/MER/SPI
Route de Trégastel - BP 40
22301 LANNION CEDEX
FRANCE

## SUMMARY

The purpose of this paper is to present for the Tahiti region, a low latitude station, a method for determining monthly, weekly and daily forecasts of radio propagation conditions.

The monthly forecasts are necessary to establish radio links in good conditions (antennas, frequencies, power and working hours determination). They are based on the forecast of the solar activity.

The weekly and daily forecasts take into account the fine evolution of both solar and magnetic activities and also the influence of ionospheric storms.

## I - INTRODUCTION

In spite of the prodigious development of telecommunications by satellite, the ionospheric links still play an important part in the radioelectrical communications. They make use of the ionospheric refraction and reflection properties. The ionosphere is not stable. It is subject to numerous variations, mainly depending of the solar activity, which conditionnally increase or decrease an ionospheric radio link signal. Thus it becomes evident that it is necessary to establish radioelectric propagation conditions forecastings.

The monthly forecasts are necessary to establish radio links in good conditions (antennas, frequencies, power and working hours determination). They are based on the forecast of the solar activity.

But the day to day variation of the critical frequency of the F2 layer (foF2), one of the most important characteristic, presents large variation in comparison of the median. Thus it is necessary to establish weekly and daily forecasts to take into account the fine evolution of both solar and magnetic activities and also the influence of ionospheric storms.

The purpose of this paper is to present for the TAHITI region, a low latitude station, a method for determining monthly, weekly and daily forecasts of radio propagation conditions.

## II - DATA BASE

The present study analyses the hourly values of the critical frequency of the main ionospheric layers (foEs, foF1, foF2) and the M3000F2 factor obtained by the ionosonde of the TAHITI station.

The ionospheric station of TAHITI (17°44S, 210°41'E) is located in the equatorial anomaly region (-17° of geomagnetic latitude). It is the only ionosonde station actually operating in the region situated in the tropical south Pacific.

Such a station is characterized by large day-time electron densities depending of the upwards ExB drifs. These lift ionization to great heights above the geomagnetic equator and then diffuse downwards to both the north and the south increasing the peak electron densities and consequently the foF2.

We research relation between these data and both solar and magnetic index.

The daily solar index R are those published by Waldmeir (Eidgenössishe sternwarte Zurich Switzerland) till 1980 and those of KOECKLLENBERGH A. (sunspot index data centre Brussels) since 1981.

The magnetic activity is characterized by the daily planetary index Ap monthly published in journal of geophysical research by the world data center A for solar-terrestrial Physics NOAA, Boulder, Colorado 80302.

A sudden variation of this index define the sudden storm commencement (SSC) which can generate sudden ionospheric disturbances (SID) or ionospheric storms.

The solar and magnetic index are defined in universal time so that to compare these data with foF2 we must work in universal time (UT).

## III - MONTHLY FORECASTING

The region, centered on TAHITI ionospheric station, for which the forecastings are established, cover approximatively a distance of 3000 km in longitude and 1000 km in latitude (fig. 1). We suppose that in such a region the ionosphere is unidimensional : the different ionospheric caracteristics are only depending of the altitude.

To characterize the ionospheric model we consider the following ionospheric characteristics :

- the monthly median ordinary frequency of the F2 layer (foF2)
- the monthly median factor of the F2 layer (M3000F2)
- the monthly 30 % ordinary frequency of the sporadic E layer foEs(30%)
- the r factor:foF2(90%)/ foF2
- the s factor:foF2(30%)/ foF2

Polynomial regression lines are then fitted between each of the above parameter in fonction of the solar activity for each month of the year and for each hours of the day :

- parabolic regression lines for the foF2 and M3000F2 characteristics
- linear regression lines for others characteristics (foEs30%, r, s).

Thus for each above parameters we have a set of 12 x 24 graphics. Some examples of such a fitting are shown in fig. 2, 3, 4.

We have now for each parameter 24 values to characterize the daily variation for different value of solar activities. A natural method to represent that variation in fonction of the hour of the day is to make a FOURIER analysis of the series. Some examples of such a fitting are shown in fig. 5 and 6.

For calculate the different characteristics of the links (Luf, Muf and the corresponding elevation angle for two probabilities 30 and 90 %), we suppose that the wave trajectory is plane and contained in the vertical plane of the link. The links cover the distance between 0 to 3000 km. The paths taking into account for calculate the characteristics of the link are described underneath.

| D(km) / Probability | $0 < D \leq 2250$ km | $2250 < D \leq 3000$ km |
|---|---|---|
| 90 % | 1F1, 1F2, 2F2, 1E | 1F1, 1F2, 2F2, 2E |
| 30 % | 1F1, 1F2, 2F2, 1E, 1Es | 1F1, 1F2, 2F2, 2E |

The 2250 km lenqth is critical because it is a limit lenqth for a normal E and a sporadic E path.

For each path we calculate the different characteristics of the link for different standard value of attenuation between a transmitter and a receiving station. The upper Muf and the lower Luf are then retained to characterize the link. Some examples of such forecasting are shown in fig. 7, 8.

## IV - WEEKLY FORECASTING

We have divided the year into 52 standard weeks which repartition is mentionned in fig. 9 with their order number in the course of the year. The February 29$^{th}$ and December 31$^{th}$ days are not taken into consideration.

Then for each standard week (w) and for each hour (h) we have determined :

1) The average weekly value of foF2 for quiet days e.g. for days for which the index Ap is less than a value equal to ten (QWfoF2 (w,h)). The average value of the solar index are also calculated for these days (QWR). The technique of prediction of the average quiet foF2 is based on the dependance of this ionospheric parameter on the mean value of the daily solar index during these days. Regression lines are fitted between the ionospheric parameter and this solar index using the following relationship :

$$QWf_0F_2(w,h) = A(w,h)X^2 + B(w,h)X + C(w,h)$$

where X is the mean value of the daily solar index QWR on quiet days. The set of the different coefficients (52 x 24) are evaluated by the least square method. The figure 10 shows one example of such a fitting. The characteristics of each point are defined by two numbers which are respectively the

year and its weight e.g. the number of days on which the average is calculated. The figure 11 shows different curves at 0 and 12 H UT at different seasons (spring (w=12), summer (w=26), autumn (w=38) and winter (w=51)). It appears that these curves present different features. Someone are increasing whith great value of solar activity, some others are decreasing.

2) The average value of foF2 (PWfoF2(w,h)) for days which Ap index is greater than ten and the corresponding average value of Ap (PWAp).

We compute the difference DWfoF2(w,h) between the two preceding values characterizing the quiet and the perturbed ionosphere on a week base.

$$DWf_0F2(w,h)=PWf_0F2(w,h)-QWf_0F2(w,h)$$

Regression lines are then fitted between DWfoF2(w,h) and PWAp index using the following relationship :

$$DWf_0F2(w,h)=D(w,h)Y+E(w,h)$$

where Y is the average value of Ap (PWAp) characterizing magnetic activity during perturbed days.

The set of the different coefficients (52 x 24) are also evaluated by the least square method. Some examples of such a fitting are shown in fig. 12 and 13.

Now to forecast the weekly hourly value of foF2 for a given week (w) we use the following relationship :

$$Wf_0F2(w,h)=A(w,h)Rw^2+B(w,h)Rw+C(w,h)+D(w,h)WAp+E(w,h)$$

where RW and WAp are respectively the predicted value of solar and magnetic activity for the week w.

Comparisons between forecastings and measured values has been made. The calculated standard deviation is equal to 1.25 MHz. The evolution of this parameter in function of day-time and year are shown in fig. 14 and 15. In this last figure we have also mentionned the evolution of the solar activity. Fig. 16 shows a histogramm of the predicted errors for the period 1971-1986.

## V - DAILY FORECASTING

The daily forecastings are relative to a day of a standard week previously defined. We have similarly proceed that for weekly forecasting.

The set of coefficients defining the evolution of foF2 in function of the solar activity is identical to this evaluated for the weekly forecasting.

To determine the influence of the daily geomagnetic index on the ionospheric characteristic, we calculated for each hour the difference between the daily values of foF2 (foF2(D,w,h)) and the mean weekly value for quiet days of foF2(QWfoF2(w,h)).

$$DDf_0F2(D,w,h)=f_0F2(D,w,h)-QWf_0F2(w,h)$$

We fitted again regression lines between this parameter and the daily geomagnetic index Ap using the following relationship :

$$DDf_0F2(D,w,h)=F(D,w,h)Z+G(D,w,h)$$

where Z is the daily geomagnetic index. The set of the different coefficients (52 x 24) are also evaluated by the least square method. Fig. 17 and 18 show different regression lines whose comments are identical to those of fig. 12 and 13.

To forecast the daily hourly values of foF2 for a given day we use the following relationship :

$$f_0F2(D,w,h)=A(w,h)R^2+B(w,h)R+C(w,h)+F(D,w,h)Ap+G(D,w,h)$$

where R and Ap are respectively the predicted daily value of solar and geomagnetic activity index for the day D.

Comparisons between forecastings and measured values have also been made for this station. The standard deviation is equal to 1.74 MHz (fig. 19 and 20). Figure 21 shows a histogram of the predicted errors for the period 1971-1986.

## VI - GEOMAGNETIC AND IONOSPHERIC PERTURBATION

We recall the following definition referring to geomagnetic and ionospheric perturbations :

- The irregular variations of geomagnetic field are consecutive to ionospheric current system alteration on the influence of a solar event. We distinguish the magnetic bay, the sudden impulse and geomagnetic storms. Among the latter we have the gradual and sudden magnetic storms. The gradual storm follows the 27 day recurrence consequently to the solar rotation and appear chiefly during low solar activity. The sudden magnetic storm is leading up by a chromospheric eruption. It result from a sudden contraction of the magnetosphere under the influence of the increased solar wind by a solar flare.

- The ionospheric storms are perturbations of the ionospheric characteristics (electron density, height of the different layers, etc...) on a period of some days. They are generally preceded and accompanied by geomagnetic storms. We have also two types of ionospheric storms : the recurrent storms specially during low solar activity with a period of 27 days corresponding to the solar rotation. The sudden commencement storm are correlated to solar chromospheric eruption and are preceded and accompanied by geomagnetic sudden storm commencement (SSC). We present hereafter two ionospheric storms corresponding to severe magnetic storms.

The $13^{th}$ july 1982 geomagnetic storm was relatively severe marked by a peak of Kp=9. Fig. 22 shows the foF2 variation at Tahiti station comparatively to the monthly value. The $13^{th}$ july we observe large variation of foF2 (6 to 14 MHz). The day after the foF2 value is about 2 MHz above the monthly values with a normal return the third day. It is a positive phase storm.

The geomagnetic storm of $5^{th}$ september 1982 (22H49UT) is also a severe storm with a peak maximum Kp=9 . This storm (fig. 23) is followed by an increase of foF2 in the afternoon. During the following night and day after foF2 values fluctuable abruptly for recovering a normal situation the following day. It is a positive phase followed by an negative phase storm.

## VII - CONCLUSION

The main advantage of this method for monthly, weekly and daily forecastings lies in its simplicity. The regression curves having been established once for all, the monthly, the weekly and daily predictions can be done by a computer in a few lapse of time.

Nevertheless these methods use mainly a fixed sample of data relative to past solar cycles and which are not renewed automatically as time elapses. Moreover the regression analysis has an averaging effect which tends to rub out any small scale variation and particularly the effect of ionospheric storms. That is particularly true on the variation of DWfoF2 and DDfoF2 in function of the geomagnetic index.

These considerations have brought us to the conclusion that ionospheric predictions would be better by means of adaptive methods. The future development in this field tend to make the prediction in a more automatic manner. They are concerned with the modelling of ionospheric disturbances and more rapid data collection particularly the measurements of geomagnetic field in this station.

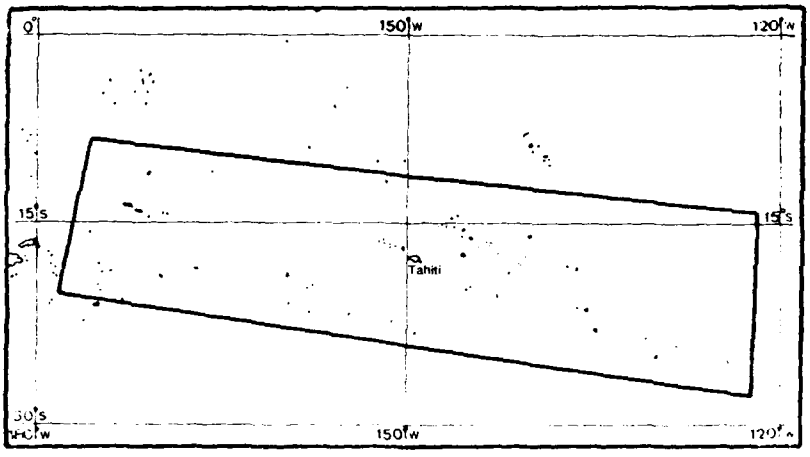



*Figure 1*

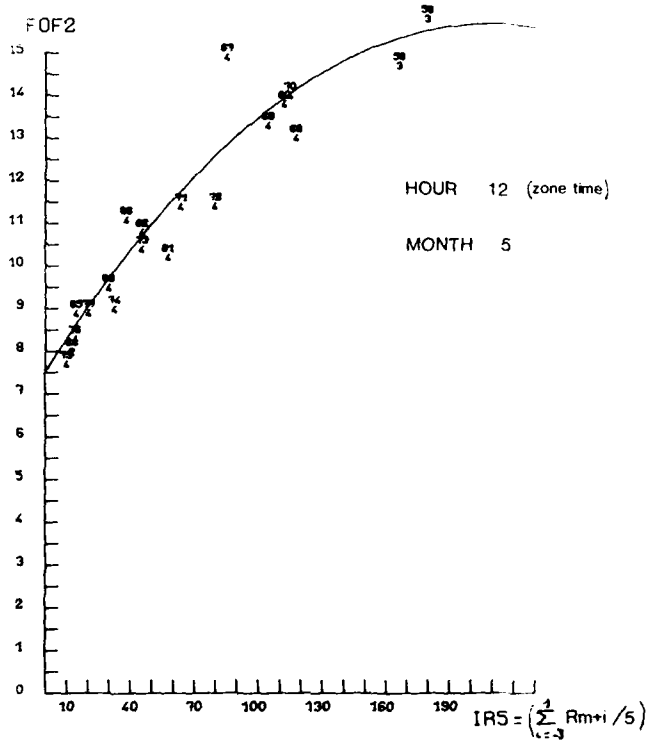



*Figure 2*

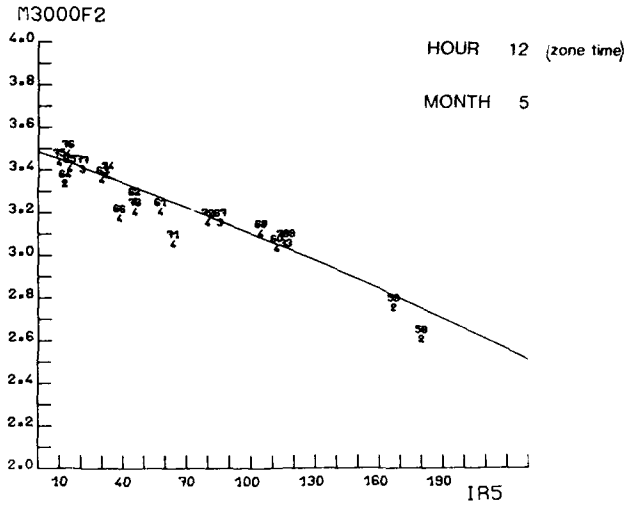



*Figure 3*

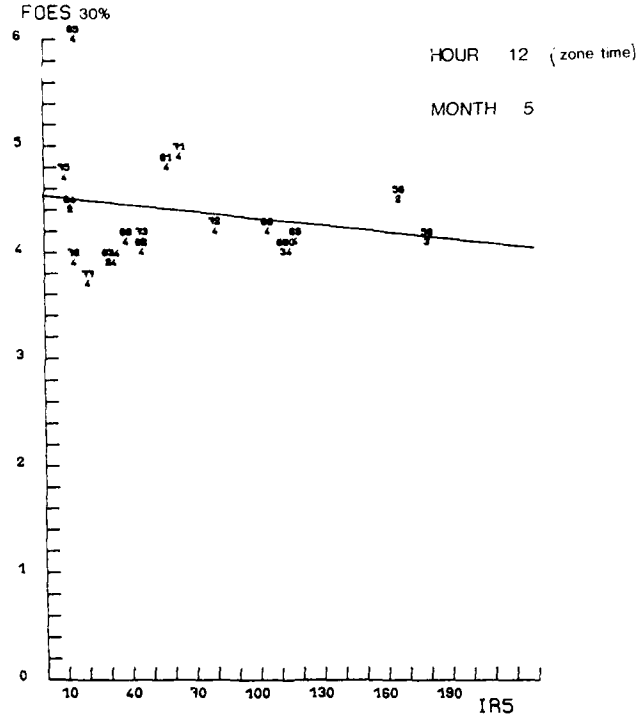



*Figure 4*

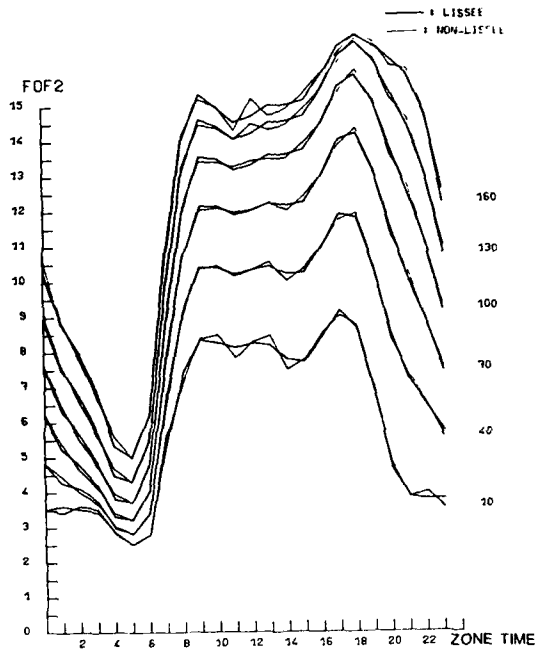



*Figure 5*

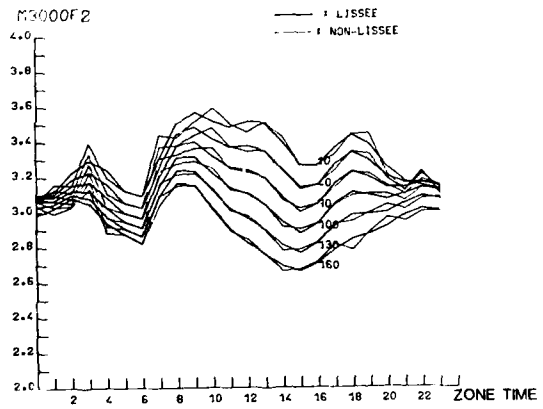
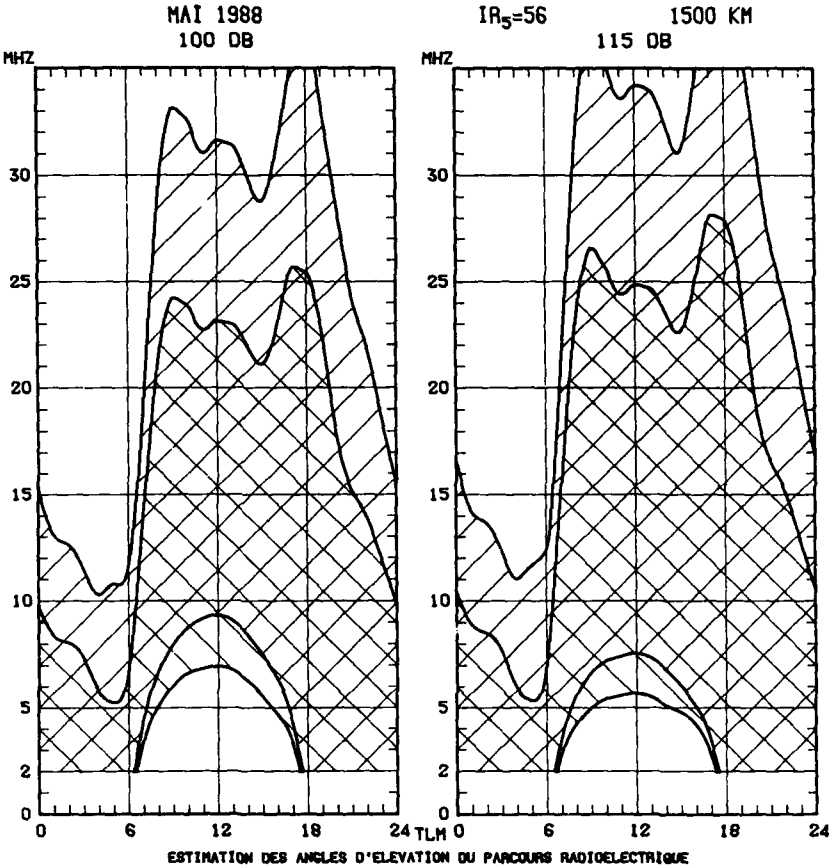


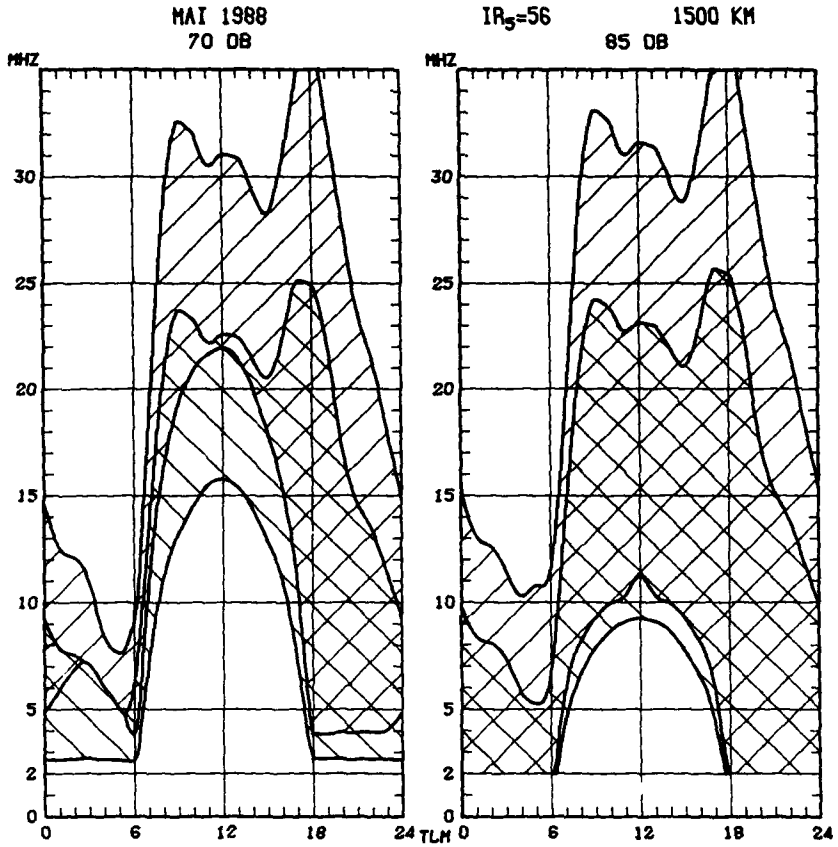

*Figure 6*




ESTIMATION DES ANGLES D'ELEVATION OU PARCOURS RADIOELECTRIQUE

| | TLM | 0 | 2 | 4 | 6 | 8 | 10 | 12 | 14 | 16 | 18 | 20 | 22 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 100 DB | MUF30 | 16 | 15 | 5 | 5 | 13 | 14 | 14 | 15 | 16 | 14 | 15 | 15 |
| | MUF90 | 16 | 15 | 16 | 17 | 13 | 14 | 14 | 15 | 16 | 14 | 15 | 15 |
| | LUF90 | | | | | 2.5 | 2.5 | 24 | 26 | 26 | | | |
| | LUF30 | | | | | 2.5 | 23 | 24 | 26 | 26 | | | |
| 115 DB | MUF30 | 16 | 15 | 5 | 5 | 13 | 14 | 14 | 15 | 16 | 14 | 15 | 15 |
| | MUF90 | 16 | 15 | 16 | 17 | 13 | 14 | 14 | 15 | 16 | 14 | 15 | 15 |
| | LUF90 | | | | | 22 | 23 | 24 | 26 | 26 | | | |
| | LUF30 | | | | | 22 | 23 | 24 | 26 | 26 | | | |

*Figure 7*




ESTIMATION DES ANGLES D'ELEVATION DU PARCOURS RADIOELECTRIQUE

| TLM | | 0 | 2 | 4 | 6 | 8 | 10 | 12 | 14 | 16 | 18 | 20 | 22 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 70 DB | MUF30 | 16 | 15 | 16 | 17 | 13 | 14 | 14 | 15 | 16 | 14 | 15 | 15 |
| | MUF90 | 16 | 15 | | 17 | 13 | 14 | 14 | 15 | 16 | 14 | 15 | 15 |
| | LUF90 | 9.5 | 8.5 | | 11 | 6.5 | 7 | 7.5 | 8.5 | 9.5 | 2.5 | 8.5 | 8.5 |
| | LUF30 | 9.5 | 8.5 | 9 | 11 | 2.5 | 2.5 | 7.5 | 8.5 | 9.5 | 2.5 | 8.5 | 8.5 |
| 85 DB | MUF30 | 16 | 15 | 5 | 5 | 13 | 14 | 14 | 15 | 16 | 14 | 15 | 15 |
| | MUF90 | 16 | 15 | 16 | 17 | 13 | 14 | 14 | 15 | 16 | 14 | 15 | 15 |
| | LUF90 | | | | | 11 | 11 | 7.5 | 11 | 11 | | | |
| | LUF30 | | | | | 2.5 | 2.5 | 2.5 | 2.5 | 26 | | | |

*Figure 8*

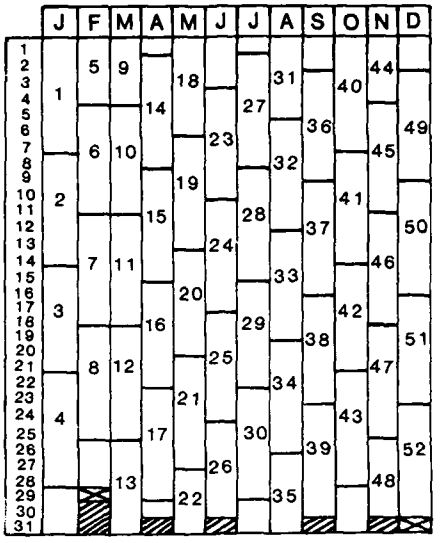



*Figure 9*

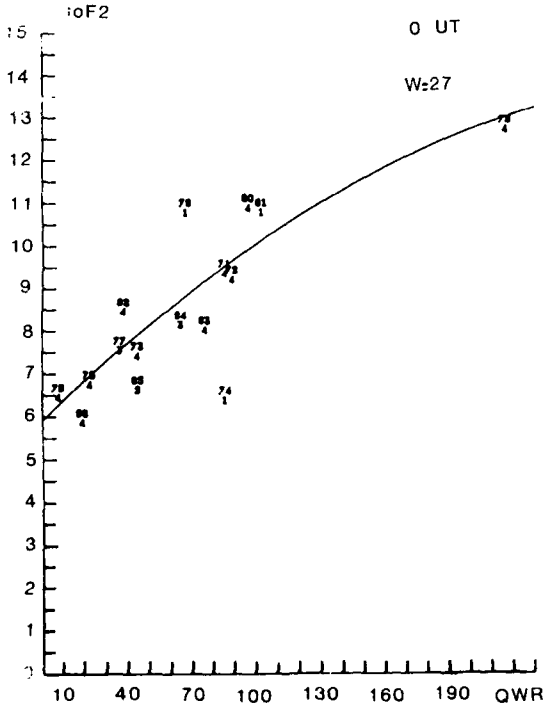



*Figure 10*

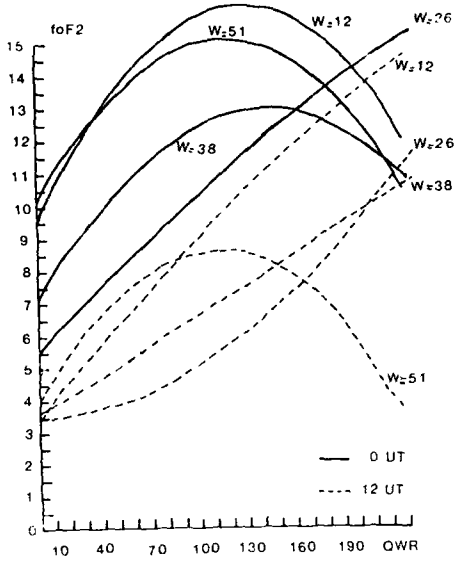



*Figure 11*

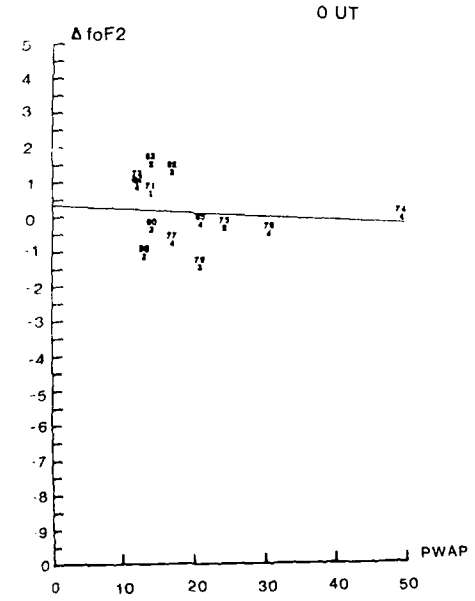



*Figure 12*

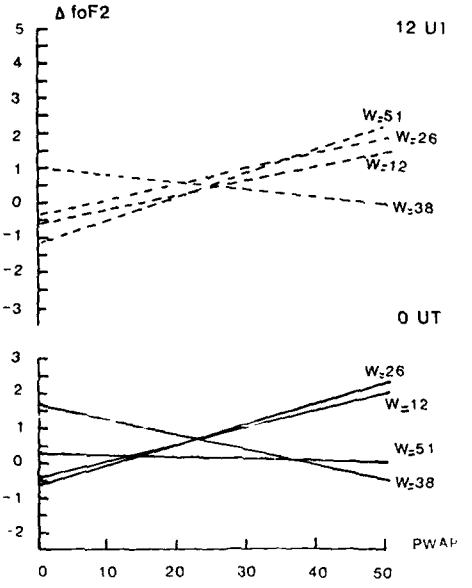



*Figure 13*

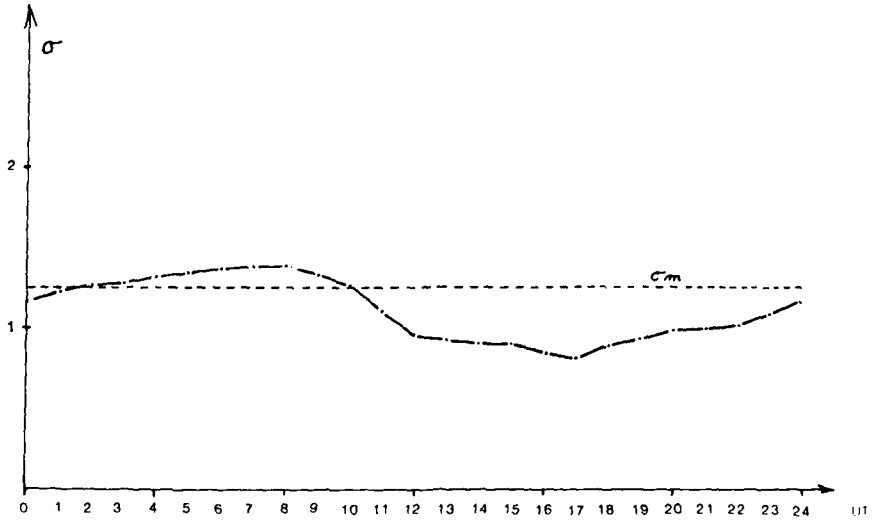



*Figure 14*

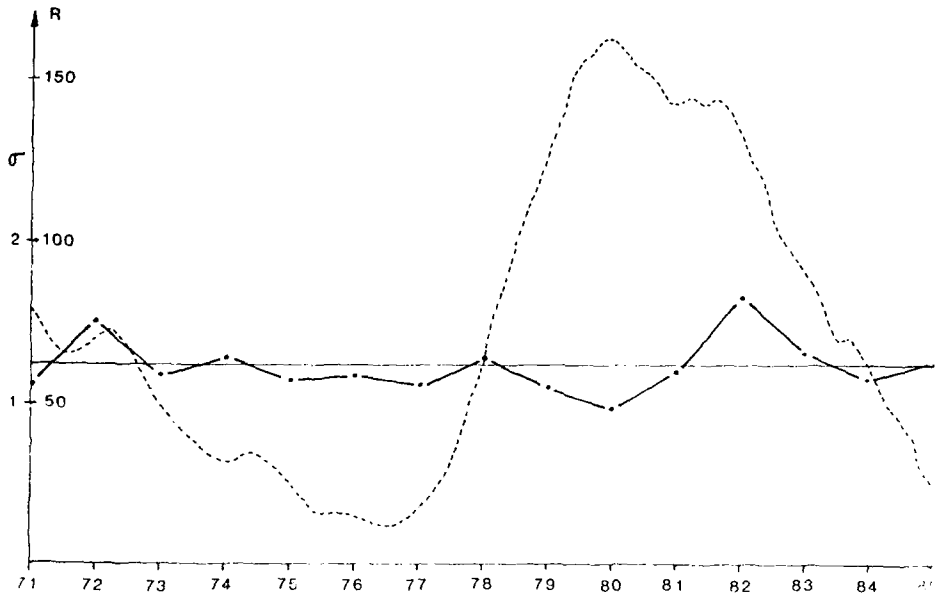



Figure 15

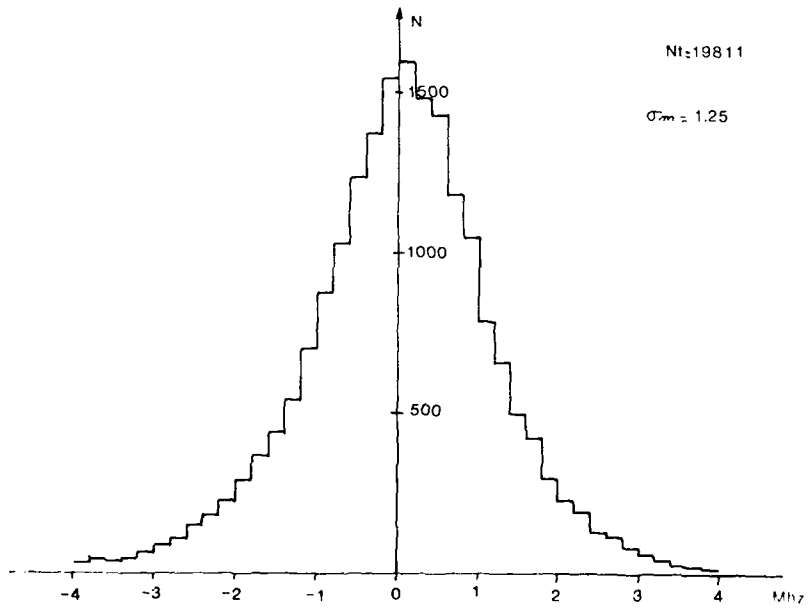



Figure 16

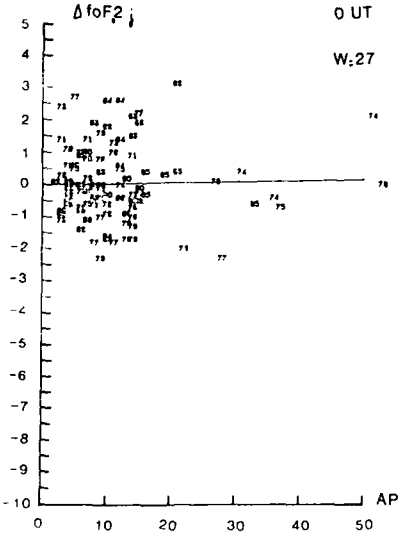



*Figure 17*

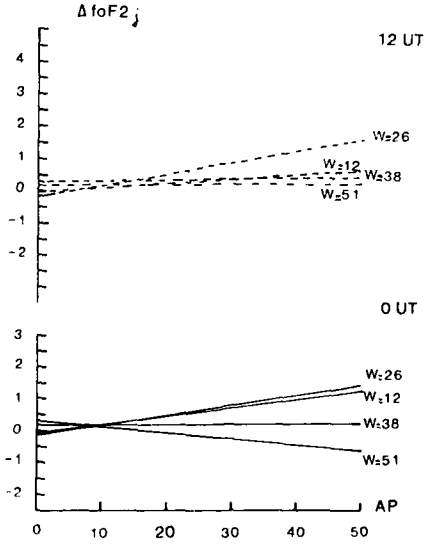



*Figure 18*

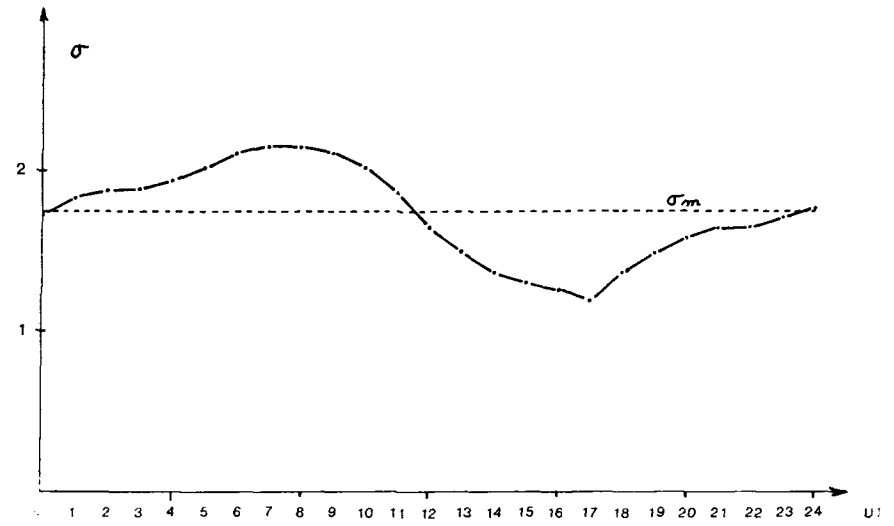



Figure 19

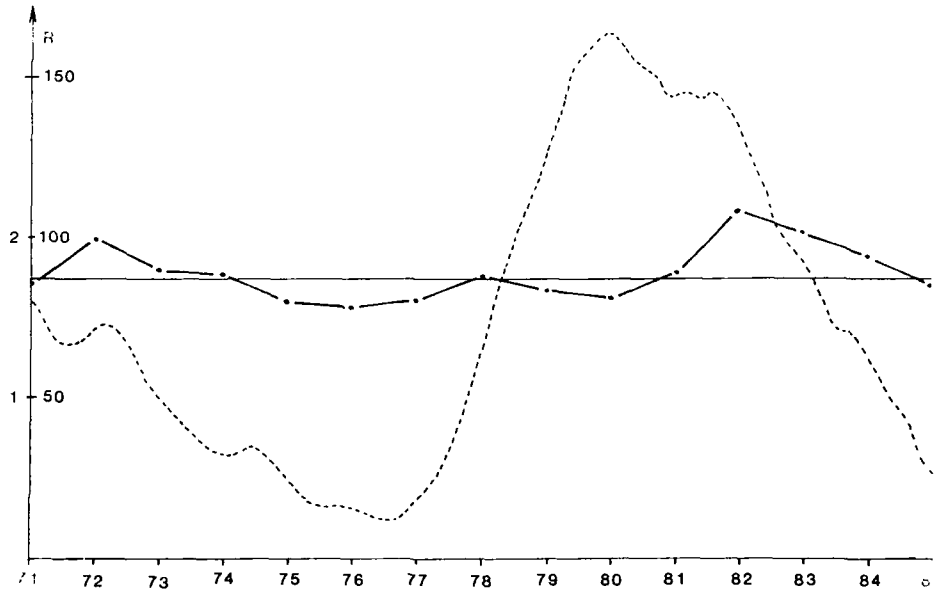



Figure 20

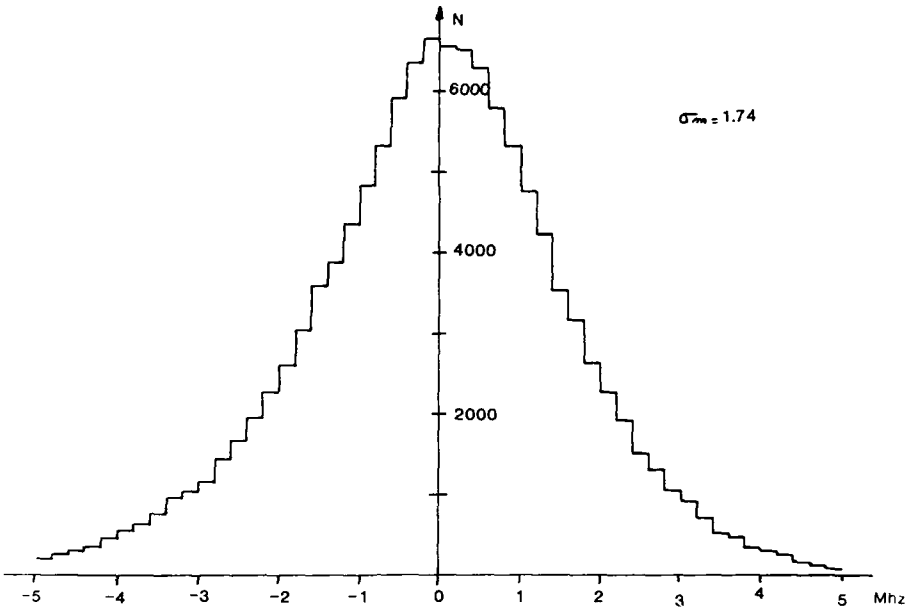



*Figure 21*

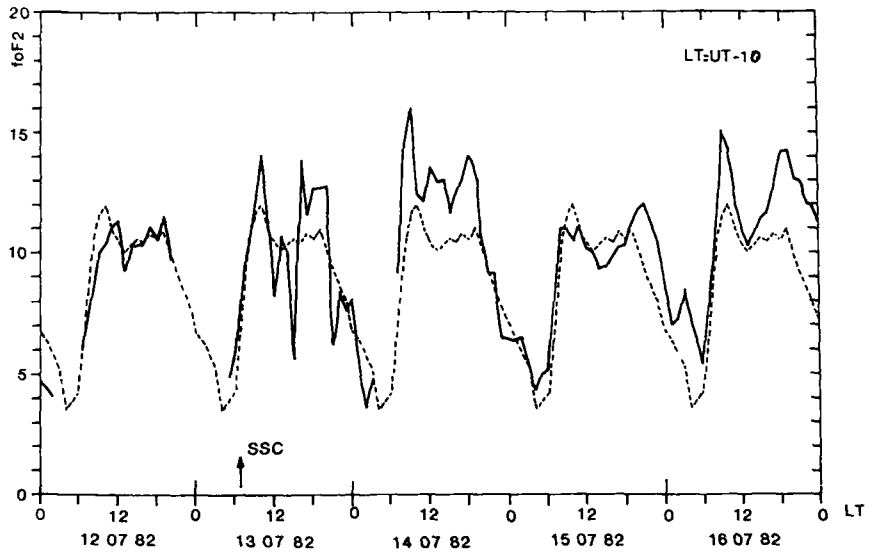



*Figure 22*

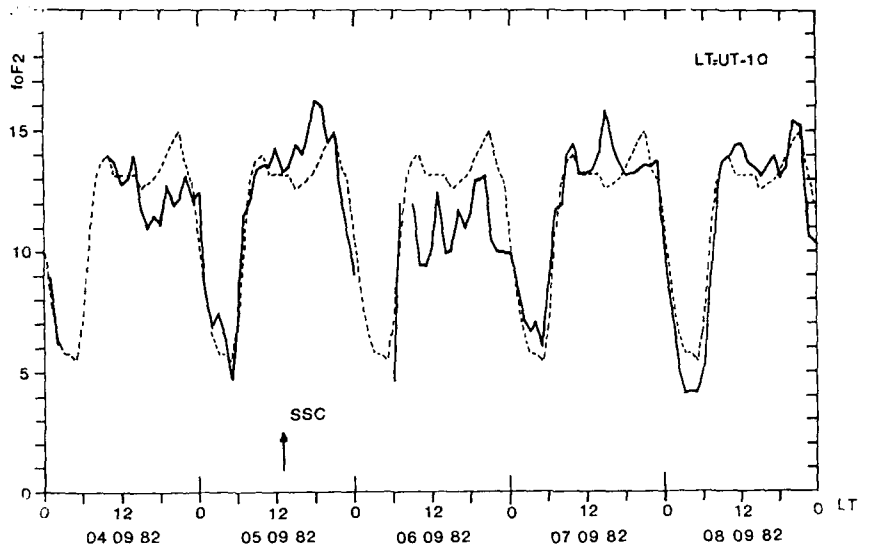



*Figure 23*

## DISCUSSION

**T.Damboldt**

1. A standard deviation of $\sigma = 1.25$ MHz or $\sigma = 1.75$ MHz for $f_0F_2$ gives the impression of a rather accurate prediction. If these values of $\sigma$ are applied to an oblique path MUF, the relevant standard deviation then becomes $\sigma = 3.75$ or $\sigma = 5.25$ MHz, respectively (assuming an M-factor of about 3). This no longer supports the impression of an "accurate prediction." Unfortunately, however, this is the present state of the art.

2. Regarding the prediction of the frequency range between LUF and MUF for the different levels of signal strength, have you taken into account propagation above the basic MUF? In other words, is the upper frequency limit the basic MUF or the operational MUF?

**Author's Reply**

The basic MUF is first calculated for each mentioned paths. We select the greater one with taking into account available operational means of intercourse. Then we apply a correcting formula to take into account of both the presence of extraordinary mode and diffusion. It is our operational MUF.

# AN IONOSPHERIC MODEL FOR HF SKY-WAVE BACKSCATTER RADAR

George H. Millman
P.O. Box 64
DeWitt, New York 13214, U.S.A.

and

Carl A. Bowser
Roger W. Swanson
Government Electronic Systems Division
General Electric Company
P.O. Box 4840
Syracuse, New York 13221, U.S.A.

## SUMMARY

The determination of target location by an HF over-the-horizon backscatter radar is accomplished by the use of an ionospheric model in conjunction with ionogram transmission curves. In this paper, an ionospheric model is described in terms of the temporal and spatial characteristics of the various ionospheric layers. The model includes the high-latitude region and is used with the transmission curve overlay process for HF backscatter radar-target coordinate registration.

## 1.0. INTRODUCTION

The geographic coordinates of a target detected by a line-of-sight radar are derived from the measurement of the elevation and azimuth angles and the time delay (i.e., radar range).

For an HF sky-wave backscatter radar, the measurement data available for target identification are the bearing angle and the time delay. The determination of the target location requires that the virtual height of reflection of the radar transmissions also be known. An estimate of the virtual reflection height can be made utilizing the concept of transmission curves overlayed on an ionogram (i.e., signal time delay versus frequency), recorded by a vertical incidence ionospheric sounder situated in the vicinity of the midpoint between the radar and the target. If a midpoint vertical incidence ionospheric sounder is not available, an ionospheric model in conjunction with a locally recorded vertical incidence ionogram can be employed to estimate the virtual height for the target coordinate registration process.

An ionospheric model applicable to the HF backscatter radar over-the-horizon identification problem is described in this paper. The procedure for radar-target coordinate determination using the ionogram transmission curve overlay approach is also presented.

## 2.0. IONOSPHERIC MODEL

### 2.1. *INTRODUCTION*

The ionospheric model defines the temporal and spatial characteristics of the solar-E and auroral-E layers, the sporadic E layer, the F1 and F2 layers, the mid-latitude trough, the auroral zone, and the auroral oval ionization. The ionospheric parameters required for model characterization are the critical frequency of the layers, the height of maximum ionization of the layers, and the semithickness of the layers. These parameters permit the use of selected quasi-parabolic layer forms in calculating virtual height trend corrections to apply to the ionogram transmission curve overlay data to achieve an estimate of the desired midpoint ionospheric virtual height.

### 2.2. *E LAYER*

#### 2.2.1. SOLAR-E LAYER

The normal solar-E layer is a region of ionization located at an altitude between approximately 90 and 130 km. The maximum ionization, which is at an altitude between 100 and 120 km, is solar controlled. That is, the ionization level increases with decreasing solar zenith angle.

The E-layer critical frequency of the ordinary wave, $f_oE$ (in MHz), is described by a modified version of the relationship by Davies (1965)

$$f_oE = 0.9\,((180 + 1.44\,SSN)\cos X_E)^{0.25} \tag{1}$$

where SSN is the sunspot number (AFGWC, 1982). The parameter $X_E$ is the normalized E-layer solar zenith angle, X (in deg), and is defined by

$$X_E = 90^\circ\,(X/101^\circ) \qquad \text{for } 0^\circ \le X \le 101^\circ. \tag{2}$$

A solar zenith angle of 101° corresponds to the zenith angle of the sun when illuminating the E layer at an altitude of 120 km and is derived on the assumption of the absence of a screening altitude (i.e., the altitude below which solar radiation is absorbed). The minimum value of $f_oE$ is set at 0.6 MHz (Air Force Global Weather Center (AFGWC), 1982).

The height of maximum ionization of the solar-E layer, $h_mE$ (in km), is given by (AFGWC, 1982)

$$h_mE = 100 + 20 \log_e (\sec X_E) \tag{3}$$

This expression which is a slightly modified version of that given by Elkins and Rush (1973) applies to the condition $X \leq 76.78^\circ$. When $X > 76.78^\circ$, the height of the E-layer maximum is maintained at 120 km.

The semithickness of the solar-E layer, $y_mE$ (in km), is defined by (Rush et al., 1982)

$$y_mE = (1/5.5)\, h_mE. \tag{4}$$

### 2.2.2. AURORAL-E LAYER

In the high latitudes, ionization is introduced into the E layer primarily by solar wind interaction with the geomagnetic field. During solar geophysical disturbances resulting in auroral particle precipitation, an E layer is formed which is present both day and night. The auroral-E ionization is in addition to the ionization of the normal solar-E layer.

The northern boundary of the auroral-E zone is taken as coincident with the center of the Feldstein-Starkov (1967) auroral oval. In other words, as shown in Figure 1,

$$\phi_N = \phi_C = (1/2)\,(\phi_E + \phi_P) \tag{5}$$

where $\phi_N$ is the corrected geomagnetic latitude of the northern boundary of the auroral-E zone, $\phi_c$ is the corrected geomagnetic latitude of the center of the auroral oval, and $\phi_E$ and $\phi_p$ are the corrected geomagnetic latitude of the equatorward edge and the poleward edge of the auroral oval, respectively. The southern boundary of the auroral-E zone, $\phi_S$ (in corrected geomagnetic latitude), is coincident with the equatorward edge of the auroral oval.

The maximum value of the auroral-E critical frequency of the ordinary wave, $(f_oE_a)_M$ (in MHz), is expressed as (AFGWC, 1982)

$$(f_oE_a)_M = 2.5 + 0.11\,Q \qquad \text{for } 0 \leq Q \leq 2.7 \tag{6}$$

$$(f_oE_a)_M = -1.0 + 1.4\,Q \qquad \text{for } 2.7 < Q \leq 4.2 \tag{7}$$

$$(f_oE_a)_M = 3.2 + 0.4\,Q \qquad \text{for } Q > 4.2 \tag{8}$$

where Q is the magnetic activity index.

The corrected geomagnetic latitude at which $(f_oE_a)_M$ exists, $\phi_M$, is given by (Rush et al., 1982)

$$\phi_M = \phi_E \qquad \text{for } Q \geq 3 \tag{9}$$

$$\phi_M = \phi_E + 1^\circ \qquad \text{for } Q < 3. \tag{10}$$

For condition $Q < 3$, the critical frequency of the auroral-E layer, $f_oE_a$ (in MHz), at the southern boundary of the auroral-E zone is related to $(f_oE_a)_M$ by (AFGWC, 1982)

$$f_oE_a = 0.6\,(f_oE_a)_M. \tag{11}$$

The value of $f_oE_a$ at the northern boundary of the auroral-E zone, $\phi_N$, is assumed to be identical to that given by Equation (11).

At a corrected geomagnetic latitude south of $\phi_M$ (between $\phi_M$ and $\phi_S$), $\phi_{MS}$, applicable for $Q < 3$, $f_oE_a$ is determined by linear interpolation according to

$$f_oE_a = 0.6\,(f_oE_a)_M + 0.4\,(f_oE_a)_M\,(\phi_{MS} - \phi_S)/(\phi_M - \phi_S). \tag{12}$$

Irrespective of the Q value, for a latitude north of $\phi_M$ (i.e., between $\phi_M$ and $\phi_N$ or $\phi_{MN}$), the critical frequency is given by

$$f_oE_a = (f_oE_a)_M - 0.4\,(f_oE_a)_M\,(\phi_{MN} - \phi_M)/(\phi_N - \phi_M). \quad (13)$$

The altitude associated with the maximum auroral-E critical frequency, $(h_mE_a)_M$ (in km), is derived from (AFGWC, 1982)

$$(h_mE_a)_M = 185 \qquad \text{for } (f_oE_a)_M < 2 \text{ MHz} \quad (14)$$

$$(h_mE_a)_M = 185 - 30\,((f_oE_a)_M - 2) \qquad \text{for } 2 \le (f_oE_a)_M < 3.5 \text{ MHz} \quad (15)$$

$$(h_mE_a)_M = 145 - 10\,((f_oE_a)_M - 3) \qquad \text{for } 3.5 \le (f_oE_a)_M \le 9 \text{ MHz} \quad (16)$$

At the outer edges of the auroral-E zone (i.e., at $\phi_S$ and $\phi_N$), the altitude of the auroral-E layer, $h_mE_a$ (in km), is given by (AFGWC, 1982)

$$h_mE_a = (h_mE_a)_M + 10. \quad (17)$$

For the condition $Q < 3$, linear interpolation if used to determine $h_mE_a$ between the latitudes $\phi_M$ and $\phi_S$. This results in

$$h_mE_a = (h_mE_a)_M + 10 - 10\,(\phi_{MS} - \phi_S)/(\phi_M - \phi_S). \quad (18)$$

For latitudes poleward of $\phi_M$, $h_mE_a$ is derived by linear interpolation according to

$$h_mE_a = (h_mE_a)_M + 10\,(\phi_{MN} - \phi_M)/(\phi_N - \phi_M). \quad (19)$$

This expression is applicable for all values of Q.

The semithickness of the auroral-E layer, $y_mE_a$ (in km), is related to the layer height by (Rush et al., 1982)

$$y_mE_a = (1/3)\,(h_mE_a) \quad (20)$$

### 2.2.3. COMBINED SOLAR-E LAYER AND AURORAL-E LAYER

When the solar-E layer and the auroral-E layer are both present, the sequence used for combining the layers is as follows (AFGWC, 1982):

For $| h_mE_a - h_mE | > 10$ km,

a) Use auroral-E layer procedure outlined in Section 2.2.2 if $(f_oE/f_oE_a) < 1.1$

b) Use solar-E layer procedure outlined in Section 2.2.1 if $(f_oE/f_oE_a) \ge 1.1$

For $| h_mE_a - h_mE | \le 10$ km,

$$(h_mE)_{AV} = (1/2)\,(h_mE_a + h_mE) \quad (21)$$

$$(f_oE)_{AV} = (1.5\,(f_oE_a)^4 + (f_oE)^4)^{1/4} \quad (22)$$

where the subscript AV indicates the combined layers. The condition $| h_mE_a - h_mE | \le 10$ km can occur when $(f_oE_a)_M$ is greater than approximately 6 MHz.

The semithickness of the combined layers, $(y_mE)_{AV}$ (in km), is specified by the function

$$(y_mE)_{AV} = (1/2)\,((1/5.5)\,h_mE + (1/3)\,h_mE_a). \quad (23)$$

### 2.2.4. SPORADIC-E LAYER

The sporadic-E layer, $E_s$ at one time called the abnormal E layer, consists of thin patches of ionization that are more intense than the normal solar-E layer in which it is embedded. The ionization can be so intense that it can prevent radio signals

from penetrating the layer. The minimum frequency at which the $E_s$ layer is no longer opaque is referred to as the $E_s$ blanketing frequency, $f_bE_s$. Above $f_bE_s$, the $E_s$ layer is partially reflecting and partially transmitting up to a maximum frequency, $f_oE_s$, beyond which $E_s$ becomes completely transparent. The parameter $f_oE_s$ is known as the critical frequency of the sporadic-E layer. The $E_s$ blanketing frequency is used in the estimation of multipath effects and, based on the analysis of Kolawole (1987), it can be assumed to be related to $f_oE_s$ by

$$f_bE_s \approx 0.7\, f_oE_s. \tag{24}$$

The characteristics of $E_s$ vary with geographic latitude. In the low latitudes, the presence of $E_s$ is a daytime phenomenon and is independent of season. In the mid-latitudes, it is prevalent during the summer daytime hours, although it can be present during the nighttime and in the winter. At high latitudes, $E_s$ occurs predominantly during the nighttime irrespective of season.

Because of the spatial and temporal statistical nature of sporadic E, it is not incorporated as a parameter in the ionospheric model. However, sporadic E is assumed to be present throughout the first-hop coverage region when it is observed by the vertical incidence (VI) ionospheric sounder at the radar site. The $f_oE_s$ in the coverage region is assigned the value of $f_oE_s$ recorded at the radar site. To conform with the sporadic-E model developed by Leftin et al. (1968), the sporadic-E layer is confined to an altitude of 110 km.

### 2.3. *F1 LAYER*

The maximum ionization of the F1 layer is concentrated in the altitude regime between approximately 180 and 200 km. The magnitude of the maximum ionization is greater than that for the normal solar-E layer. The F1 layer is solar controlled, similar to the E layer, and thus is present only during the daylight hours. It is more effective in the summer months compared to the wintertime. Its ionization level increases with increasing sunspot number and, during ionospheric-magnetic storm conditions, it can be greater than that for the F2 layer.

The critical frequency of the ordinary wave of the F1 layer, $f_oF1$ (in MHz), is represented by the function (AFGWC, 1982), which is a slightly modified version of the one defined by Davies (1965)

$$f_oF1 = (4.3 + 0.01\,SSN)\cos^n(X_{F1}) \tag{25}$$

where $X_{F1}$ is the normalized F1-layer solar zenith angle given by

$$X_{F1} = 90^\circ\,(X/105.5^\circ) \qquad \text{for } 0^\circ \le X \le 105.5^\circ. \tag{26}$$

The angle, 105.5°, corresponds to the zenith angle of the sun, X (in deg), when illuminating the F1 layer at an altitude of approximately 240 km. It is derived by assuming the absence of a screening altitude.

The exponent n, in Equation (25), is solar-zenith-angle dependent according to (AFGWC, 1982)

$$n = 0.2 \qquad \text{for } 0^\circ \le X \le 90^\circ \tag{27}$$

$$n = 0.2 + 0.3\,(X - 90^\circ)/15.5^\circ \qquad \text{for } 90^\circ < X \le 105.5^\circ. \tag{28}$$

Equation (28) shows that, for the condition $90^\circ < X \le 105.5^\circ$, the value of n varies linearly between 0.2 and 0.5. When the calculated $f_oF1$ value at a location is greater than or equal to the $f_oF2$, the F1 layer is removed from the model output.

The height of maximum ionization of the F1 layer is given by (Rush et al., 1982)

$$h_mF1 = 165 + 0.6428\,X \tag{29}$$

where $h_mF1$ is in km and X is in deg.

The semithickness of the F1 layer, $y_mF1$ (in km), is assumed to be related to $h_mF1$ (in km) by the function (Rush et al., 1982)

$$y_mF1 = 0.25\,h_mF1. \tag{30}$$

### 2.4. *F2 LAYER*

In contrast to the E and F1 layers, the F2 layer is basically not solar-zenith-angle controlled. The characteristics of the F2 layer are influenced to a great extent by solar-geophysical conditions and are a function of geomagnetic coordinates. The F2-layer ionization maximizes at an altitude on the order of 300 km. The median values of the F2-layer critical frequency, $f_oF2$, are generated by the numerical set of coefficients developed by Jones and Obitts (1970). The coefficients which are a function of geographic coordinates, universal time-of-day, month, and solar activity (i.e., sunspot number) are identical to those presently used in 1) the Ionospheric Prediction HFMUFES-4 Program developed by the Institute for Telecommunication

Sciences (Hayden et al., 1976), and 2) the AFGWC Polar Ionospheric Model (1982). The $f_oF2$ value obtained using the Jones and Obitts (1970) coefficients applies to any day of a given month.

The height of the maximum ionization of the F2 layer, $h_mF2$, is determined using the method proposed by Dudeney (1983). According to Dudeney, $h_mF2$ (in km) is given by

$$h_mF2 = [1490/(M(3000)F2 + \Delta M)] - 176 \quad (31)$$

where M(3000)F2 is the factor used for converting vertical incidence frequencies to oblique frequencies for propagation to a ground distance of 3000 km via the F2 layer. The parameter, $\Delta M$, is the correction to the M(3000)F2 factor and is expressed by

$$\Delta M = [0.253/(x-1.215)] - 0.012 \quad (32)$$

where

$$x = f_oF2/f_oE \text{ or } 1.5, \text{ whichever is larger.} \quad (33)$$

Monthly median values of the M(3000)F2 factor, which is used in the computation of $h_mF2$, Equation (31), are specified in terms of numerical coefficients (Jones et al., 1966) and are available in the Ionospheric Communications Analysis and Prediction (IONCAP) Program developed by the Institute for Telecommunication Sciences (Teters et al., 1983).

The semithickness of the F2 layer, $y_mF2$, is derived from the ratio of $h_mF2$ to $y_mF2$ given in the IONCAP Program. The computation of $y_mF2$ is made once $h_mF2$ is determined. The numerical coefficients of M(3000)F2 and the ratio $h_mF2/y_mF2$ are derived using the same data base and apply to the fifteenth day of a given month. The monthly median values of the above parameters were derived from ionospheric sounding data taken during the 5-year period extending from 1954 through 1958.

### 2.5. *TROUGH - F2-LAYER MODIFICATION*

The trough, which is often referred to as the mid-latitude trough, the high-latitude trough, or the Ottawa trough, is a region of low electron density in the F region of the ionosphere. The geographic location of the trough, which is considered to be a nighttime phenomenon, varies between 35° N and 70° N latitude. In terms of corrected geomagnetic coordinates, the trough mostly is located between 50° N and 70° N latitude. The exact location of the trough depends on such factors as the magnetic activity, time of day, and season of the year.

The poleward wall of the trough is bounded by the auroral oval (Feldstein and Starkov, 1967) where the electron densities are higher than in the trough. In other words, as depicted in Figure 1,

$$\phi_{Nt} = \phi_E \quad (34)$$

where $\phi_{Nt}$ is the corrected geomagnetic latitude of the northern edge of the trough.

The modifications made to the F2-layer parameters in the region of the trough follow somewhat the approach used in the AFGWC Polar Ionospheric Model (1982). In the region of the trough (i.e., the region south of the equatorward edge of the auroral oval ($\phi_E > \phi$)), the critical frequency of the F2 layer is modified in accordance with the function

$$(f_oF2)_t = f_oF2\,(1 + \Delta N) \quad (35)$$

where $\phi$ is the corrected geomagnetic latitude (in deg) and

$$\Delta N = \tau\,(1 + \cos\,(2\pi(D + 11)/365)) \quad (36)$$

where D is the Julian Day.

The parameter, $\tau$, is defined by the function

$$\tau = \tau_1 \exp\,((X_1 - X_1^2)/2) \exp\,(-(t - 3)^2/12) \quad (37)$$

where

$$X_1 = (\phi - \phi_E)/X_A \quad (38)$$

$$X_A = 3.7 + 1.3\,K_p \tag{39}$$

and where $X_A$ is in deg, t is the corrected geomagnetic local time (in h) and $K_p$ is the magnetic index. When (t-3) > 12, (t-3) is replaced with (24 - (t-3)) in Equation (37).

The parameter, $\tau_1$, in Equation (37) is solar zenith angle and time dependent in accordance with the following conditions:

$$\tau_1 = 0 \qquad \text{for } X \le 90^\circ \text{ or } 6 < t < 18 \text{ h} \tag{40}$$

$$\tau_1 = -0.2 \qquad \text{for } X > 105.5^\circ \text{ and } 18 \le t \le 24 \text{ h and } 0 \le t \le 6 \text{ h} \tag{41}$$

$$\tau_1 = -0.2\,(X - 90^\circ)/15.5^\circ \qquad \text{for } 90^\circ \le X \le 105.5^\circ \text{ and } 18 \le t \le 24 \text{ h and } 0 \le t \le 6 \text{ h.} \tag{42}$$

The maximum height of the F2-layer trough occurs at a corrected geomagnetic latitude, $\phi_{Mt}$ (in deg), defined by

$$\phi_{Mt} = \phi_E - 1.5^\circ. \tag{43}$$

At this latitude, the maximum F2-layer trough height, $(h_mF2)_{Mt}$ (in km), which exists at nighttime, is given by

$$(h_mF2)_{Mt} = 450 - 100\,V \tag{44}$$

where

$$V = (t-3)/4 \qquad \text{for } 3 \le t \le 6 \text{ or } 18 \le t \le 21 \text{ h} \tag{45}$$

$$V = -(t-3)/6 \qquad \text{for } 0 < t < 3 \text{ h} \tag{46}$$

$$V = -(t-27)/6 \qquad \text{for } 21 < t \le 24 \text{ h.} \tag{47}$$

The maximum value of V is limited to 1 for calculated values of V greater than 1. In addition, for $(h_mF2)_{Mt} < h_mF2$, $(h_mF2)_{Mt}$ is assumed equal to $h_mF2$.

Linear interpolation is used to determine the height of the F2 layer within the trough between the latitudes $\phi_{Mt}$ and $\phi_E$ and between $\phi_{Mt}$ and $\phi_{St}$, the southern boundary of the trough. The value of the height of the F2-layer ionization peak at the southern boundary of the trough is determined according to Section 2.4 and by Section 2.6 at the northern boundary.

The corrected geomagnetic latitude of the trough's southern boundary is determined according to the following criteria:

$$\phi_{St} = \phi_E + (3.7 + 1.3\,K_p)\,(0.5 - \sqrt{R}) \qquad \text{for } \tau_1 < 0 \text{ and } R > 0.25 \tag{48}$$

$$\phi_{St} = \phi_E - 1.5^\circ \qquad \text{for } \tau_1 < 0 \text{ and } R \le 0.25 \tag{49}$$

$$\phi_{St} = \phi_E - 1.5^\circ \qquad \text{for } \tau_1 \ge 0 \tag{50}$$

where

$$R = 0.25 - ((t-3)^2/6) + 2\log_e(-10\tau_1) + 2\log_e(1 + \cos(2\pi(D + 11)/365)). \tag{51}$$

The semithickness of the F2-layer trough is derived using the same approach applied to the normal F2 layer (Section 2.4).

### 2.6. *AURORAL ZONE - F2 LAYER MODIFICATION*

The auroral zone, shown in Figure 1, is considered the region encompassed within the boundary of the auroral oval as specified by Feldstein and Starkov (1967). It is a region of enhanced ionization with respect to the normal-E layer and F layer. The model of $f_oF2$ in the auroral zone is similar to that employed in the AFGWC Polar Ionospheric Mode (1982).

For a location in the auroral zone satisfying the condition, $\phi_C \ge \phi > \phi_E$, the critical frequency of the auroral zone F2 layer, $(f_oF2)_A$, is given by

$$(f_oF2)_A = ((f_oF2)^2 + 9\Delta N)^{1/2} \tag{52}$$

where

$$\Delta N = \exp(-X_2^2/2) \tag{53}$$

and where

$$X_2 = (\phi_E - \phi)/(\phi_P - \phi_E). \tag{54}$$

For a location north of the center of the auroral oval, $\phi > \phi_C$, the parameter, $\Delta N$, is modified to the form

$$N = \exp(-2X_2^2). \tag{55}$$

In the auroral zone and in the polar cap (i.e., $\phi > \phi_P$), the value of the maximum height of the F2 layer, $(h_mF2)_A$ (in km), is defined by

$$(h_mF2)_A = (1490/M(3000)F2) - 206. \tag{56}$$

The retardation imposed by the E layer and the F1 layer and identified in Equation (31) is replaced by a constant 30 km. For the auroral zone, $(h_mF2)_A$ is independent of geomagnetic coordinates and time. The M(3000)F2 factor in Equation (56) is identical to that contained in the IONCAP Program (Teters et al., 1983). The semithickness of the F2 layer in the auroral zone is derived in a similar manner to that used of the normal F2 layer (Section 2.4).

### 2.7. *AURORAL OVAL*

The model of the auroral oval (belt) incorporated in the HF radar ionospheric model is the one developed by Feldstein and Starkov (1967) from photographic observations of aurorae made during the International Geophysical Year (IGY), 1957-1958, a period of intense solar activity. The Feldstein-Starkov auroral oval defines the condition of auroral activity in the high latitude ionosphere.

The location of the equatorward edge of the auroral oval is represented by the function (Starkov, 1969)

$$\phi_E = 72 - 0.9\,Q - 5.1\cos(15\,(t-0.8)). \tag{57}$$

This expression is applicable for $Q \geq 1$. The coordinates of the equatorward edge of the auroral oval for Q equal to 0, given in Table 1, were scaled from the data plots in the Feldstein-Starkov (1967) publication. In order to avoid discontinuities from one time-block to another, the value of $\phi_E$ is applied to the middle of the time-block which is referred to as the Average Time in Table 1. Linear interpolation is used to determine $\phi_E$ at other times and intermediate nonlinear Q values.

The coordinates of the poleward edge of the auroral oval, presented in Table 2, were also sccaled from the Feldstein-Starkov (1967) data. The corrected geomagnetic latitudes are assumed to apply to the middle of the time block, similar to Table 1. Linear interpolation is used to deduce the location of the poleward edge of the oval for other times and Q values. For negative Q values, Q is assumed to be 0 and, for Q greater than 7, Q is assumed to be 7.

### 2.8. *MAGNETIC INDICES, $K_P$ AND Q, EQUIVALENCE*

The southern and northern boundaries of the auroral oval, as presented by Feldstein and Starkov (1967), are derived as a function of the magnetic indices, $K_p$ and Q, for the time block 22-02 h, corrected geomagnetic local time. From an analysis of the data, it is found that the functional relationship between these two magnetic indices can be expressed by (AFGWC, 1982)

$$Q = 3\,K_p \qquad \text{for } K_p < 1 \tag{58}$$

$$Q = K_p + 2 \qquad \text{for } K_p \geq 1 \tag{59}$$

This approach for the conversion of the two parameters has been adopted for the ionospheric model.

The magnetic activity index, $K_p$, is a mean 3-hour magnetic reading derived from observations made between 47° and 63° geomagnetic latitude. The scale of 0 to 9 is an arbitrary one where 0 refers to very quiet conditions and 9 to extremely disturbed magnetic conditions (Lincoln, 1967).

The magnetic activity index, Q (often referred to as the polar range index), is a quarter-hourly index which is, at times, used to indicate the severity of magnetic conditions for high latitudes, above about 58° geomagnetic latitude. The index has not been widely employed in geophysical analysis because of the complexity involved in data reduction of magnetic observations to obtain the parameter (Lincoln, 1967).

The AFGWC provides to its product users an effective Q value, $Q_e$, derived from energetic electron-satellite measurements of the corrected geomagnetic latitude of the equatorward boundary of the auroral oval (Gussenhoven et al., 1981). When the satellite data are not readily available, $Q_e$ can be estimated from Equations (58) and (59).

In addition, AFGWC derives on a daily basis an estimate of the effective sunspot number, $SSn_e$, deduced from VI soundings. Both $Q_e$ and $SSN_e$ are used in place of Q and SSN in the ionospheric model.

### 2.9. *REAL TIME MODEL CORRECTION*

The ionospheric model provides a median estimation of the temporal and spatial characteristics of the various ionospheric layer parameters on a global basis. The median estimates of $f_oE$ and $f_oF1$ have been found to be within 2 to 5 percent of the actual daily values for the E- and F1-layer variables and $f_oF2$ estimates within 10 to 25 percent of the actual F2-layer critical frequency (Rush and Gibbs, 1973).

To minimize the day-to-day variability, data from VI ionospheric sounders are used to adjust the ionospheric model over zones in the vicinity of the VI sounder sites. Linear correction is applied to the E and F1 variables. In the case of the F2 layer, model adjustment is made utilizing the correlation coeffients of the F2-layer parameters derived by Rush and Miller (1973) in conjunction with the ionospheric world map generation method developed by Zacharisen (1965).

When VI data are not available on a given day due to, for example, an equipment malfunction, 5-day average values are substituted from the days just prior to the data loss as the best estimator of the missing data for adjusting the ionospheric model (Rush and Fibbs, 1973).

Figures 2, 3, and 4 are plots of the spatial characteristics of $f_oF2$, $h_mF2$ and the MUF(3000)F2, respectively, computed using the ionospheric model. The parameter, MUF(3000)F2, is the maximum usable frequency for propagation to 3000-km distance via the F2 layer.

## 3.0. RADAR-TARGET COORDINATE DETERMINATION

### 3.1. *INTRODUCTION*

The determination of the signal path between an HF transmitter and receiver has been and continues to be of interest to all users of this radio frequency band. The capability for detailed analytic modeling of the ionospheric plasma and for calculation of the Poynting vector energy propagation paths has been developed to make such determination possible. Additionally, more approximate ray path modeling of the radio signal/propagation interactions has been developed for rapid VI ionospheric sounding record interpretation. For over-the-horizon backscatter (OTH-B) radar application, ray path approximations (i.,e., transmission curve overlays) are combined with analytic quasi-parabolic layer modeling to provide propagation path specification to support radar coordinate conversion. The transmission curve overlay solution, developed by Smith (1939) for VI ionogram interpretation, has been extended to oblique sounding - HF radar applications.

The single quasi-parabolic layer modeling of the ionosphere has found broad applications since ray path integrals can be evaluated exactly in closed form (Croft and Hoogasian, 1968). Exact ray paths of a multiple quasi-parabolic layer ionosphere have been shown to be readily solvable using minicomputers and microcomputers (Hill, 1979) and can aid in real-time propagation assessment systems.

The ray path solutions for virtual height calculated from ionospheric model parameters are the means of extrapolating transmission curve overlay solutions to areas widely removed from the radar VI sounding site. For propagation conditions where a VI site is close to the ray path reflection zone, the transmission curve overlay solutions are used directly without application of the ionospheric model. The ionospheric model supports the analysis of multipath propagation (when deemed necessary by assessment of vertical incidence soundings) by validating supportability of alternate modes of propagation (multiple hops or multiple layer paths.)

### 3.2. *IONOSPHERIC SOUNDER AT PATH MIDPOINT*

The provision of midpath VI sites on oblique propagation paths permits direct application of transmission curve overlays as a means of obtaining virtual heights of reflection for propagation to specific ground ranges at different OTH-B radar operating frequencies.

The VI data are processed in accordance with the automated VI sounding systems developed by Reinisch and Huang Xuequin (1983). The ordinary ray VI traces are used solely in the propagation path calculations.

A VI record and transmission curve overlay are illustrated in Figure 5 where the transmission ground distances are labeled for each 400 nmi from the radar site and the elevation angle curves are labeled at each 8 deg. The ionospheric trace shown in Figure 5 supports propagation only to 1800 nmi at the overlay operating frequency of 22.5 MHz. The application of transmission curve overlays has been automated to generate only low-angle (i.e., non-Pedersen rays) solutions from the ionogram trace. From the transmission curve overlay, it is possible to identify all multiple layer solutions to the same ground range as well as multiple hops from the same layer. Propagation modes found in the transmission curve overlay processing which are not currently identified for radar conversion use are identified to radar operators to ensure correct multipath labeling of radar echoes.

The procedure employed for the generation of the transmission curve is presented in the Appendix. From the virtual height and elevation angle information obtained using the transmission curve overlay, a coordinate registration table can be computed for the target location process. The overlay data are compared with precalculated tables to determine the possible solutions for each trace height on the selected radar operating frequency. The distance solutions are restricted to low-angle

conditions by initially performing a localization of the skip distance (i.e., minimum ground range obtained by overlay solution on the frequency) for the radar operating frequency and then incrementing in ground range out to the maximum ground range supported by the layer.

### 3.3. *IONOSPHERIC SOUNDER AT RADAR SITE*

In the absence of midpath VI data, an overhead VI sounding at the radar site is used to provide base transmission curve overlay solutions that are extrapolated to the midpoint using the ionospheric model. An overhead VI ionogram is illustrated in Figure 6 where the ordinary ray layer trace (curve labeled P-F2) was entered by operator action. The automated overlay solutions are applied as in the midpath sounding case. Gradients present in the ionospheric model due to the radar site and midpath separation are taken into account in the calculations. The virtual height calculated from the modeled ionospheric parameters accounts for the diurnal variability of the layers by automated selection of the true height quasi-parabolic forms appropriate to the number of layers and electron density magnitude.

Figure 7 presents the general form of the electron density profiles provided for the extrapolation calculations:

a) Case 1 - two layers separated by an intermediate electron free region;

b) Case 2 - two layers with an intermediate region of constant electron density;

c) Case 3 - three layers with discontinuous electron density slope at each layer's critical frequency. The layers are quasi-parabolic functions with the semithickness adjusted to ensure the junction of two layers at the height of maximum density of the underlying layer.

The analytical approach of Croft and Hoogasian (1968) is employed in the virtual height calculations for a single layer. When multiple layers are present, Hill's (1979) analytical functions are exercised. The virtual height calculations at the VI radar site and the midpath range provide a measure of the virtual height offset that must be applied to the midpath transmission curve overlay solutions. In this approach, the percentage offset (indicated by the midpath calculated virtual height to VI radar site calculated virtual height ratio) is used as an expansion or contraction factor to apply to the radar VI derived transmission curve overlay solutions.

It is noteworthy that this approach does not expand the frequency interval of overlay solutions from those of the VI radar site and thus does not totally account for the largest density variations between the VI radar site and the path midpoint that occur during sunrise and sunset conditions.

## 4.0. CONCLUSIONS

The application of a real-time corrected ionospheric model to support the over-the-horizon backscatter radar coordinate conversion process results in a means of accounting for measured temporal and spatial variations from median predictive conditions.

The combination of radar site vertical incidence ionospheric soundings with transmission curve overlays gives a usefull approximation of the virtual height of reflection encountered by an HF backscatter radar.

## 5.0. REFERENCES

AFGWC Polar Ionospheric Model, Air Force Global Weather Central, Program Listing, February 1982.

Croft, T.A. and H. Hoogasian, "Exact Ray Calculations in a Quasi-Parabolic Ionosphere with No Magnetic Field," Radio Science, Vol. 3, pp. 69-74, January 1968.

Davies, K., "Ionospheric Radio Propagation," National Bureau of Standards, Monograph 80, U.S. Government Printing Office, April 1965.

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," Journal of Atmospheric and Terrestrial Physics, Vol. 45, pp. 629-640, 1983.

Elkins, T.J. and C.M. Rush, "A Statistical Predictive Model of the Polar Ionosphere," in an Empirical Model of the Polar Ionosphere, Air Force Cambridge Research Laboratories, AFCRL-TR-73-0331, Air Force Surveys in Geophysics, No. 267, May 1973.

Feldstein, Y.I. and G.V. Starkov, "Dynamics of Auroral Belt and Polar Geomagnetic Disturbances," Plantetary Space Science, Vol. 15, pp. 209-229, 1967.

Gussenhoven, M.S., D.A. Hardy and W.J. Burke, "DMSP/F2 Electron Observations of Equatorward Auroral Boundaries and Their Relationship to Magnetospheric Electric Fields," Journal of Geophysical Research, Vol. 86, pp. 768-778, February 1, 1981.

Haydon, G.W., M. Leftin and R.K. Rosich, "Predicting the Performance of High Frequency Sky-Wave Telecommunication Systems (The Use of the HFMUFES-4 Program)," Institute for Telecommunication Sciences, U.S. Department of Commerce, OT Report 76-102, September 1976.

Hill, J.R., "Exact Ray Paths in a Multisegment Quasi-Parabolic Ionosphere," Radio Science, Vol. 14, pp. 855-861, September-October 1979.

Jones, W.B., R.P. Graham and M. Leftin, "Advances in Ionospheric Mapping by Numerical Methods," National Bureau of Standards, U.S. Department of Commerce, NBS Technical Note 337, May 1969 (Reissued as ESSA Technical Note ERL 107-ITS 75, May 1969).

Jones, W.B. and D.L. Obitts, "Global Representation of Annual and Solar Cycle Variation of $f_0F2$ Monthly Median 1954-1958," Institute for Telecommunication Sciences, U.S. Department of Commerce, OT/ITS Research Report 3, October 1970.

Kolawole, L.B., "The Transparency Characteristics of $E_s$ Types," Radio Science, Vol. 13, pp. 159-165, January-February 1978.

Leftin, M., S.M. Ostrow and C. Preston, "Numerical Maps of $f_oE_s$ for Solar Cycle Minimum and Maximum," Institute for Telecommunication Sciences, U.S. Department of Commerce, ESSA Technical Report ERL 73-ITS 63, May 1968.

Lincoln, J.V., "Geomagnetic Indices," in Physics of Geomagnetic Phenomena, International Geophysics Series, Vol. 1, pp. 67-100, Academic Press, New York, 1967.

Reinisch, B.W. and Huang Xuequin, "Automatic Calculation of Electron Density Profiles from Digital Ionograms, 3. Processing of Bottomside Ionograms," Radio Science, Vol. 18, pp.l 477-492, May-June 1983.

Rush, C.M. and J. Gibbs, "Predicting the Day-to-Day Variability of the Mid-Latitude for Application to HF Propagation Predictions," Air Force Cambridge Research Laboratories, AFCRL-TR-73-0335, Air Force Surveys in Geophysics, No. 268, May 1973.

Rush, C.M. and D. Miller, "A Three-Dimensional Ionospheric Model using Observed Ionospheric Parameters," Air Force Cambridge Research Laboratories, AFCRL-TR-73-0567, Environmental Research Paper No. 455, September 1973.

Rush, C.M., R.K. Rosich, C.B. Brooks, D.L. Leise and M. PoKemper, "A Simplified Model of the High Latitude Ionosphere for Telecommunications Applications," Institute for Telecommunication Sciences, U.S. Department of Commerce, NTIA Report 82-94, January 1982.

Smith, N., "The Relation of Radio Sky-Wave Transmission to Ionosphere Measurements," Proceedings IRE, Vol. 27, pp. 332-347, May 1939.

Starkov, G.V., "Analytical Representation of the Equatorial Boundary of the Oval Auroral Zone," Geomagnetism and Aeronomy (English Edition), Vol. 9, pp. 614, 1969.

Teters, L.R., J.L. Lloyd, G.W. Haydon and D.L. Lucas, "Estimating the Performance of Telecommunication Systems Using the Ionospheric Transmission Channel, Ionospheric Communications Analysis and Prediction Program User's Manual," Institute for Telecommunication Sciences, U.S. Department of Commerce, NTIA Report 83-127, July 1983.

Wieder, B., "Some Results of a Sweep-Frequency Propagation Experiment Over a 110 km East-West Path," Journal of Geophysical Research, Vol. 60, pp. 395-409, 1955.

Zacharisen, D.H., "Space-Time Correlation Coefficients for Use in Short-Term Ionospheric Predictions," National Bureau of Standards, U.S. Department of Commerce, NBS Report 8811, 1965.

## 6.0. APPENDIX - TRANSMISSION CURVE OVERLAY GENERATION

### 6.1. *VIRTUAL HEIGHT DETERMINATION*

The procedure employed for the calculation of the transmission curve overlay for virtual height determination is as follows:

a) For a given oblique transmission frequency, $f_{ob}$ (in MHz), and a given radar-target ground distance, D (in nmi), and for discrete values of virtual height, h' (in km), compute the angle of incidence at the virtual reflection height $\alpha$ (in deg), defined by

$$\alpha = \tan^{-1}\left(\sin(1.852D/2R_o)/(1 + (h'/R_o) - \cos(1.852D/2R_o))\right) \quad (60)$$

where $R_o$ is the radius of the earth (6371 km).

b) For the given value of D, calculate the correction factor, K, expressed by

$$K = A_o + A_1 D + A_2 D^2 + A_3 D^3 \quad (61)$$

where $A_o = 9.56 \times 10^{-1}$, $A_1 = 1.44 \times 10^{-4}$ (nmi)$^{-1}$, $A_2 = -6.15 \times 10^{-8}$ (nmi)$^{-2}$, and $A_3 = 2.04 \times 10^{-11}$ (nmi)$^{-3}$.

The above coefficients were computed from Wieder's (1955) correction factor-distance curve. The value of K which applies to a curved ionosphere varies between approximately 1.0 and 1.2. For a plane ionosphere, K reduces to unity.

c) For the various values of $\alpha$ derived on the basis of discrete values of h' and for the corresponding K value, determine the equivalent vertical incidence frequency, $f_v$ (in MHz), given by

$$f_v = (f_{ob})(\cos \alpha)/K. \quad (62)$$

This expression is often referred to as the corrected Secant law (Davies, 1965). The least squares curve fitted through the h' versus $f_v$ data points is the D ground distance - transmission curve to be overlayed on the VI ionogram.

d) Repeat steps a through c for ground distances $D_1$, $D_2$, $D_3 \ldots D_i$.

e) Repeat steps a through d for oblique transmission frequencies $(f_{ob})_1$, $(f_{ob})_2$, $(f_{ob})_3 \ldots (f_{ob})_j$.

### 6.2. *ELEVATION ANGLE DETERMINATION*

The procedure employed for the calculation of the transmission curve overlay for elevation angle determination is as follows:

a) For a given transmission elevation angle, E (in deg), and for the same oblique transmission frequency ($f_{ob}$) and discrete values of virtual height (h') used above, compute the ionospheric incidence angle, $\alpha$ (in deg), defined by:

$$\alpha = \sin^{-1}((R_o/(R_o + h'))\cos E) \quad (63)$$

b) For the given value of E and the derived values of $\alpha$, compute the radar-target ground distance, D (in nmi), from the expression

$$D = (R_o/1.852)(1 - (\alpha + E)/90). \quad (64)$$

c) For the various values of D, compute the correction factor, K.

d) Compute the equivalent vertical frequency ($f_v$) using the coresponding values of $\alpha$ and K determined in steps a. and c., respectively. The curve fitted through the h′ versus $f_v$ data points is the E elevation angle-transmission curve to be overlayed on the VI ionogram.

e) Repeat steps a. through d. for elevation angles $E_1$, $E_2$, $E_3 \ldots E_k$.

f) Repeat steps a. through e. for oblique transmission frequencies $(f_{ob})_1$, $(f_{ob})_2$, $(f_{ob})_3 \ldots (f_{ob})_j$.

### 6.3. *VIRTUAL HEIGHT ADJUSTMENT DUE TO TRANSMITTER-RECEIVER SEPARATION*

The procedures discussed above are applicable to both monostatic and bistatic VI sounder systems. However, in the case of a bistatic VI sounder system (i.e., transmit and receive antennas are separated by a significant distance), an adjustment in the virtual height of reflection must be made in order to take into account the increase in the apparent virtual height resulting from the transmitter-receiver separation.

In a bistatic VI sounder system, the actual (adjusted) virtual height is related to the measured virtual height, $h'_m$, by the function

$$h' = h'_m \cos\gamma + R_o(\cos(S/2R_o))-1) \quad (65)$$

where S is the separation of the transmit and receive sounder antennas (in km) and $\gamma$ is the angle between the ray path and the vertical at the path midpoint given by

$$\gamma = \sin^{-1}((R_o/h'_m) \sin (S/(2R_o)). \quad (66)$$

It can be shown that, with increasing virtual height of reflection, the cos $\gamma$ term increases to unity and the difference between the actual and measured virtual heights diminishes to zero.

A bistatic VI sounder system also requires an adjustment in the equivalent vertical frequency. The measured transmitted vertical frequency, $f_{vm}$, is slightly higher than the actual (adjusted) equivalent vertical frequency, $f_v$, according to the relationship

$$f_v = f_{vm} \cos \gamma. \quad (67)$$

The above relationships are applied in adjusting the transmission curve overlay solutions for the distortion introduced in the vertical incidence ionospheric sounding data by a bistatic VI sounder system configuration.

**ACKNOWLEDGEMENT**

The work reported in this paper was sponsored by the Department of the Air Force, Headquarters Electronic Systems Division (AFSC), under Contract F19628-82-C-0114.

TABLE 1

CORRECTED GEOMAGNETIC LATITUDE OF THE EQUATORWARD EDGE OF THE FELDSTEIN-STARKOV AURORAL OVAL FOR Q = 0

| Corrected Geomagnetic Local Time (h) | | Corrected Geomagnetic Latitude (deg) |
|---|---|---|
| Time Block | Average Time | |
| 22-02 | 00 | 70.2 |
| 02-06 | 04 | 69.6 |
| 06-10 | 08 | 73.1 |
| 10-14 | 12 | 75.9 |
| 14-18 | 16 | 75.5 |
| 18-22 | 20 | 69.2 |

TABLE 2

CORRECTED GEOMAGNETIC LATITUDE OF THE POLEWARD EDGE OF THE FELDSTEIN-STARKOV AURORAL OVAL

| Magnetic Index (Q) | Corrected Geomagnetic Latitude (deg) | | | | | |
|---|---|---|---|---|---|---|
| | Average Corrected Geomagnetic Local Time (h) | | | | | |
| | 00 | 04 | 08 | 12 | 16 | 20 |
| 0 | 71.8 | 74.2 | 74.2 | 77.6 | 77.0 | 71.6 |
| 1 | 69.8 | 73.8 | 74.3 | 76.9 | 76.4 | 71.2 |
| 2 | 70.1 | 73.5 | 74.4 | 76.2 | 75.7 | 70.8 |
| 3 | 70.8 | 73.5 | 74.6 | 75.9 | 74.9 | 71.1 |
| 4 | 71.6 | 73.6 | 74.2 | 75.6 | 74.0 | 71.3 |
| 5 | 72.2 | 73.6 | 73.5 | 75.4 | 73.4 | 71.5 |
| 6 | 72.8 | 73.6 | 72.3 | 75.2 | 72.4 | 70.9 |
| 7 | 73.5 | 73.2 | 72.0 | 75.0 | 71.6 | 70.7 |

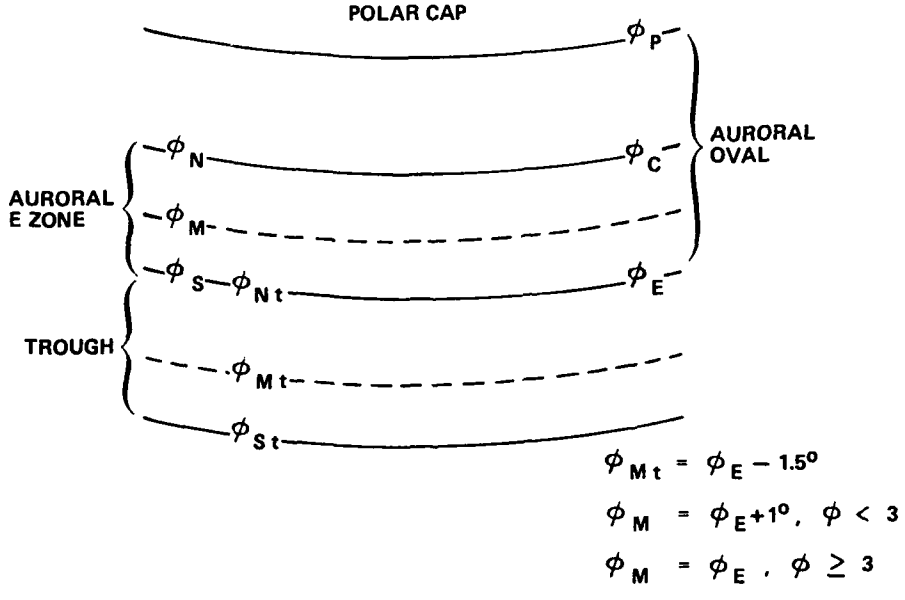



Figure 1. Geomagnetic Orientation of the High Latitude Ionosphere

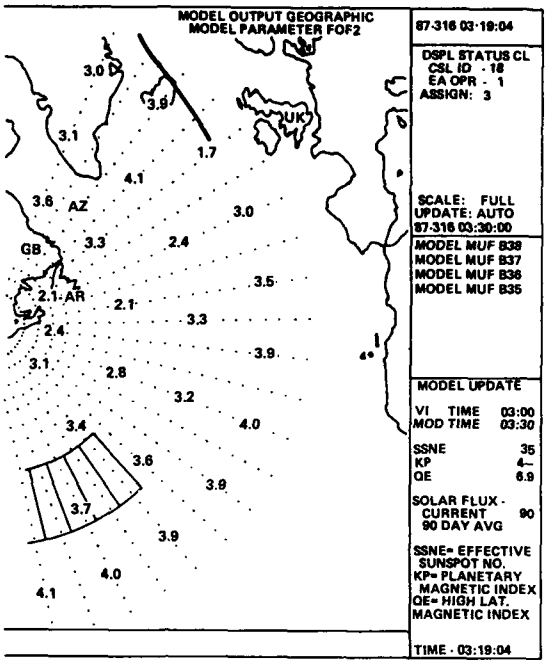



Figure 2. Spatial Characteristics of $f_oF2$ Computed by the Ionospheric Model

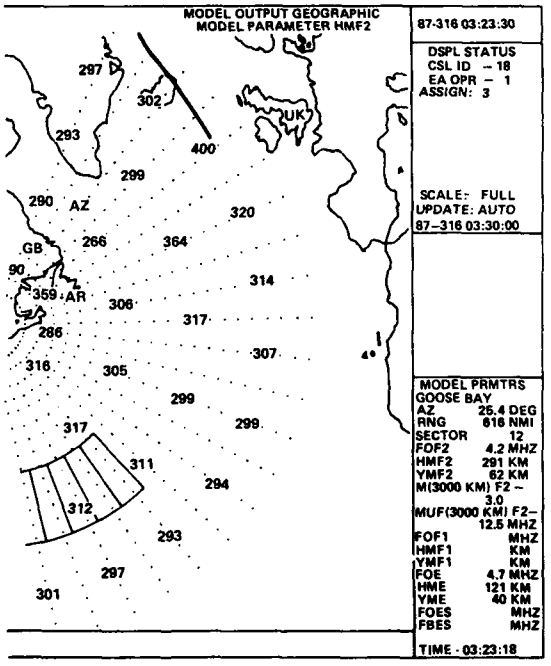



Figure 3. Spatial Characteristics of $h_mF2$ Computed by the Ionospheric Model

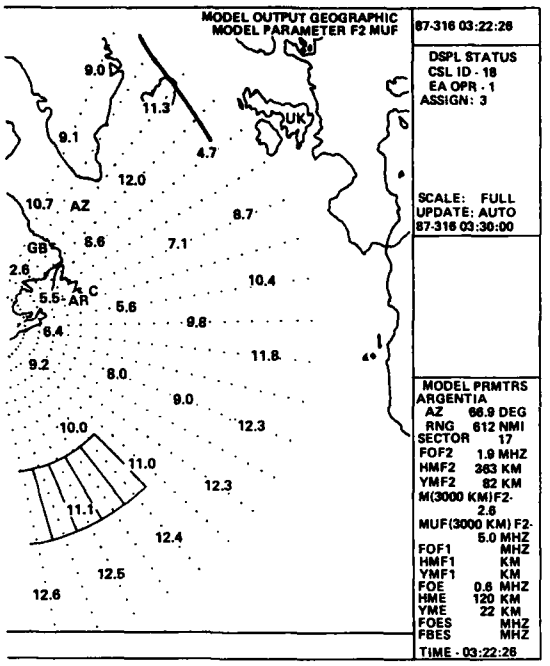



Figure 4. Spatial Characteristics of MUF(3000)F2 Computed by the Ionospheric Model

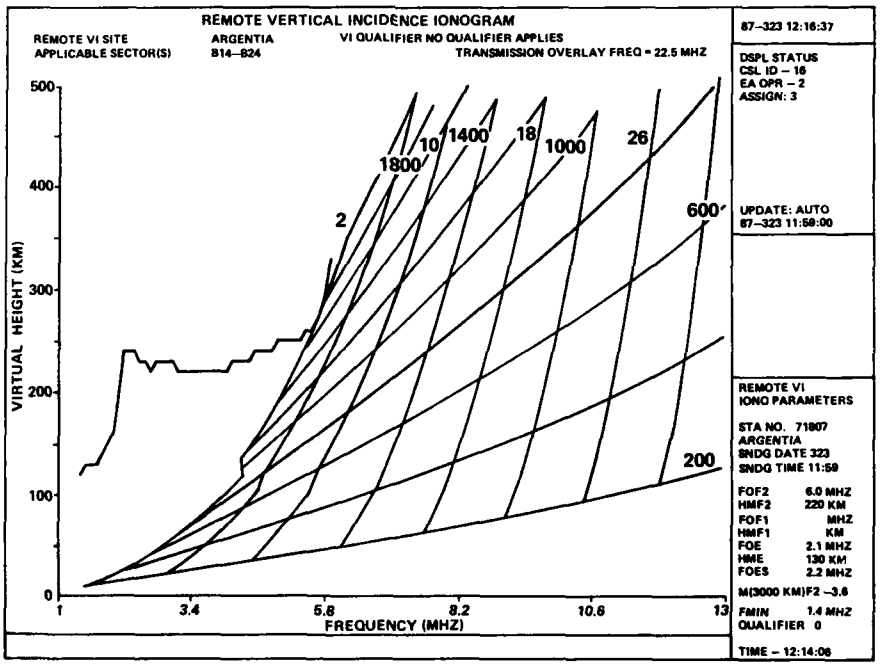



Figure 5. Scaled Remote Vertical Incidence Ionogram with Transmission Curve Overlay for an Operating Frequency of 22.5 MHz

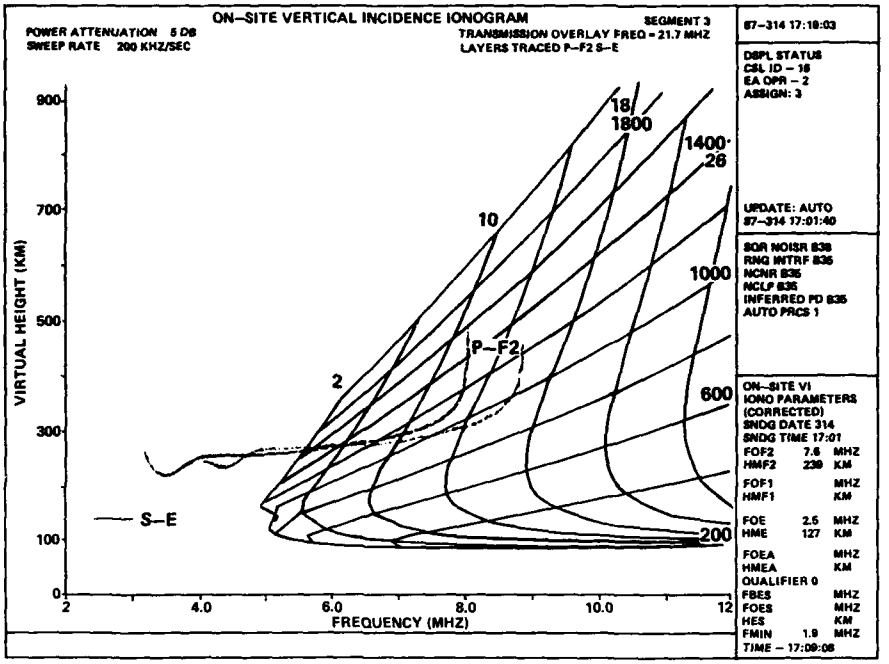



Figure 6. Vertical Incidence Ionogram Recorded at Radar Site with Transmission Curve Overlay

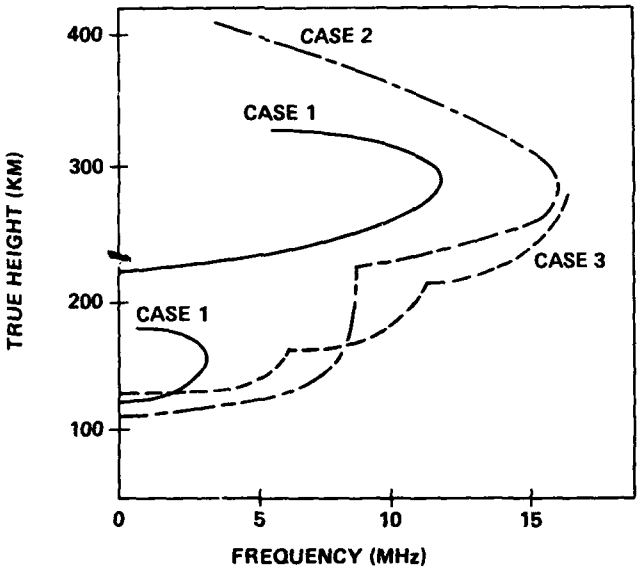



Figure 7. Electron Density Profiles Used in the Virtual Height Calculation Process

## DISCUSSION

**G.Rostoker**

I note that you have a model auroral oval with input parameters of poleward and equatorward borders. How are the predictions of those border positions achieved?

*After Answer*

You should note that the poleward border of the oval can sweep over several degrees of latitude several times in a three-hour interval (characteristic of a $K$ index).

**Author's Reply**

The Air Force supplies all effective $Q$ index based on the DMSP satellite-optical measurements of the auroral oval; the northern boundary of the auroral oval, given in Table 2, was scaled from Feldstein-Starkov data. If all effective $Q$ are not supplied, the parameter is derived from $Kp$ given by equations (58) & (59). The Air Force also furnishes a daily effective sunspot number derived from vertical incidence sounding data.

**C.Goutelard**

For your calculations we use curves of N. Smith. These curves are constructs for spherical stratification. You use sophisticated models, and we have *Horizontal Gradients Ionization*. Have you an evaluation of the error introduced by this method?

**Author's Reply**

We have not analyzed the effects of gradients on the transmission curves. We do take into account the error introduced in the transmission curve calculation by the system being bistatic and not monostatic. The adjustments to the virtual height and the equivalent vertical frequency are given by equations (65)-(67). The coefficients in the definition of the correction factor, $K$, (equation (61), were deduced from Wieder's curve. The accuracy of specifying $K$ is high.

# ACCURATE MEASUREMENTS OF THE TOTAL ANGLE OF ARRIVAL OF HF SKYWAVES

by

Z.R.Jeffrey, P.T.Middleton and C.Winkler
Andrews Antennas
Technology Park 5095
Adelaide
Australia

## ABSTRACT

This paper describes the procedures adopted in the ANDREW Single Station Location (SSL) system, SKYLOC, to measure the azimuth and elevation angle of a skywave from an HF transmitter, with a view to determining the location of that transmitter by appropriate simulation of the ionosphere and the signals propagating through it. The procedures adopted are capable of providing very accurate measurements of the azimuth and elevation angle, well within the variability imposed by the ionosphere, and are able to process signals with very small signal-to-noise ratios.

## 1 INTRODUCTION

The ANDREW Single Station Location (SSL) system, SKYLOC [Trademark], is a member of a class of systems known generically as direction finding (DF) Systems. DF systems can be divided into a number of different classes. In some, only the horizontal component of the arrival angle is measured, while in others both the horizontal and vertical components are obtained. Where only the azimuth is measured, two or more spaced stations are required to determine the position of the transmitter. Systems that measure both components are called SSL systems because, in the case of signals propagated by the ionosphere, the elevation angle and azimuth from just one station are sufficient to define the location of the transmitter, assuming a knowledge of the ionosphere along the circuit.

Historically, SSL was developed because of the problems of audibility with the traditional DF network, since it often happened that a transmitter could be heard at only one site. Also, if the range to that site was less than a few hundred km, most of the DF systems then in use could not give accurate bearings for such high elevation-angle signals. Most of the early work in this field was carried out in Australia, where an operational SSL system was fielded in the late fifties.

The simplest direction finder consists of a movable directional antenna and a receiver. The arrival angle is measured by moving the antenna until the the signal is either a maximum or a minimum. As simple antennas generally give broad maxima and sharp minima, the null method is usually used, even though it may not give good performance on weak signals. The simple loop antenna or some derivative of it appears in many direction finders.

One of the disadvantages of the null method is that the signal being measured is reduced in the null while noise and signals from other directions are not, effectively broadening the null. To overcome this problem, amplitude comparison methods, which use two directional antennas with different orientations, were introduced.

As it is not always feasible or desirable to move the antenna itself, many methods use a fixed antenna, the elements of which act as probes, and the movement occurs inside the equipment. One of the earliest systems used fixed crossed loops to feed a small orthogonal loop arrangement with a rotating loop inside it. This was called a goniometer. (Nowadays the term "goniometer" means almost any cyclic selection equipment and in fact may not be mechanical.)

The cathode-ray direction finder used crossed loops, but instead of employing a goniometer, each loop was fed to one channel of a twin-channel receiver. The outputs from the receiver were applied to the orthogonal pairs of deflection plates in a special cathode-ray tube. A clean signal produces a straight line, the angle of which gives the azimuth of the received signal.

Until the end of World War II the most common type of DF system was the crossed loop or Adcock. The Adcock system used loops from which the top horizontal member was removed and the bottom horizontal member well shielded in order to reduce polarization errors. These systems are small in relation to the wavelength of the received signals and consequently are classified as narrow aperture direction finding (NADF) systems. NADF systems are susceptible to wavefront distortions caused by re-radiation from objects near the site and wave interference.

During the war the Wullenweber system heralded the era of the wide aperture direction finding (WADF) systems. The Wullenweber has a circularly disposed antenna array (CDAA) up to 1000m in diameter with a large number of elements. About a third of the elements are combined to form sum and/or difference beams. These beams are effectively rotated in azimuth by a rotating switch (goniometer) which connects and combines elements in turn around the ring.

One further type of system, called pseudo-doppler, depends on the modulation imposed on the signal if an antenna is moved in a circle, the phase of the modulation being a function of the direction of the signal. In practice a CDAA is used and the receiver is rapidly switched to each antenna in turn around the ring.

Another class of direction finders is the so-called interferometric systems, in which the direction of arrival of the incoming wave is not deduced by rotating beams (or nulls) but from the phase measurements made on a number of spaced antennas.

Unlike the beam-forming type of WADF's, interferometers accept all the signals incident on the array. Two different approaches have been used to process the results. One, called Wave Front Analysis (WFA), attempts to recover the major rays by analysing the complex voltages appearing on the elements of the antenna array under wave interference conditions. The other, called Wave Front Testing (WFT), processes results only during times of quasi uni-modal propagation (QUMP). QUMP is detected by looking for a linear phase shift across the array and near-equal amplitudes on all the elements.

Finally, there are the time-difference direction finders (TDDF) which measure the differences between the times of arrival of the radio wave at a number of receiving sites. A hyperbolic position line is obtained from the time difference between each pair of sites.

## 2 PROCEDURES USED IN SKYLOC

A block diagram of the SKYLOC system is given in Figure 1. SKYLOC is an interferometric system, with typically seven antennas spaced along two orthogonal arms, as illustrated in Figure 2. Phase and amplitude are measured at each antenna. The inner antenna pairs in the array are separated by 10 metres. Because the wavelength at 15 MHz is 20 metres, there can be no phase ambiguities for signals which are less than 15 MHz, and we can therefore adjust the measured phase on the outer elements by the required number of two pi radians to give the correct value. For frequencies greater than 15 MHz, we use the fact that the antenna spacing of the intermediate element is a non-integral multiple of the inner antenna spacing. This allows identification and correction of the phase ambiguities. The phase measurements are subjected to various quality filters and then used to determine the elevation angle and azimuth of the incoming wavefront, as illustrated in Figure 3.

Signals received from two spaced antennas are mixed down to a centre frequency of 3kHz in two receiver chains, depicted in Figure 1, whose gain and phase characteristics are known. The receiver transfer functions are depicted as H1(w) and H2(w) respectively. If the input signals are s1(t) and s2(t), then the output power spectra are filtered by H1(w) and H2(w) and become S11(w) and S22(w). See Figure 4. They represent estimates for the power distribution with frequency, and are calculated by multiplying the Fourier transform of the output signal by its own complex conjugate. The cross spectra, on the other hand, S12(w) and S21(w), relate the power in one channel with power in the other. While power spectra are real, the cross spectra are in general complex. The magnitude |S12(w)| is a complicated expression but the phase is simple to interpret. The phase of S12(w) represents the phase difference between the signals s1(t) and s2(t), modified only by the phase difference between the receiver transfer functions. If we can calibrate these transfer functions we can make H1(w) approximately equal to H2(w) and remove the effects of different H1(w) and H2(w).

In SKYLOC, we use a 512-point transform and a bin resolution of 20 Hz. The data are windowed, and processed to give a cross-phase estimate and a quality estimate based on the distribution of the phases. Data of poor quality are rejected.

Tests are also performed to determine the departure of the wavefront from linearity, using the phase and amplitude measurements made at the intermediate "wavefront test" antennas (35 metres from the reference antenna). If this departure is sufficiently small the assumption of QUMP is made, and the data passed to the next processing stage.

The algorithms used to process these signals and construct the angles of arrival rely on a sequence of operations which excise bad data points and coherently process accepted data. This provides a signal processing gain in excess of the processing gain from the FFT alone (which would provide some 20dB or so), and results in the ability to make angle of arrival measurements on signals with a S/N ratio as low as -20dB in the HF band. Correlation processing cannot produce reliable estimate at less than about +10dB SNR.

An additional advantage of this approach is that processing may be performed on a section of the receiver bandwith, giving the ability to perform narrow-band DF operations. This enables separation of signals which occur within the same 3kHz receiver bandwidth, but which have energy components occupying different spectral bins.

## 3 TYPICAL SKYLOC ANGLE OF ARRIVAL OBSERVATIONS

The philosophy adopted in SKYLOC is that a small sample of the good quality data is much preferable to a large sample of poor quality data. Data are therefore screened very rigorously, and under some conditions it is possible for almost all of the data to be rejected.

Figure 5 shows a set of individual angle of arrival measurements made on the Melbourne-Adelaide circuit during the 1986 trials of SKYLOC [1,2] on a series of over 200 uncooperative transmitters. The equipment configuration used at the time relied on measurements with a two-channel HF receiver, which made the system slower than the present four-channel system. A full frame of angle of arrival measurements was made every second, and grouped into 19-second groups. The "cuts" column in Figure 5 shows how many of the measurements survived the filtering processes. This count was highest for those sections of

the record marked E (mode) or F1, and least for those sections corresponding to a mode transition. For example, the count was only 4 at 1.2934UT when the E mode faded, to be replaced by the F2 mode. The azimuth column gives a weighted mean of the individual azimuth estimates.

Figure 6(a) shows the full 20-minute accumulation of the elevation angle data, with peaks picked automatically at 18 degrees (1E mode), 34.6 degrees (1F1) and 45.3 degrees (1F2). Figure 6(b) shows the accumulated azimuth histograms with a peak at 121.7 degrees, a little lower than the correct value of 122.4 degrees. Each 19-second group of data is also characterized by the standard deviation of the measured azimuths, and these standard deviations are plotted in Figure 6(c).

The peak-picking procedure had three steps:-

1. The histogram values are filtered using a Parks-McClellan finite impulse response filter of length 39. The passband cut-off is >6.25 and the stopband cutoff is >12.5, with stopband attenuation of 60dB. This type of filter has the same ripple magnitude in the passband and stopband, and has linear phase response. Linear phase response means that there is no shift in the peaks.

2. A noise floor is set equal to the average value taken across all bins, with the proviso that the floor is not below a fixed value (2.0) and is not less than a fixed proportion (0.25) of the highest peak.

3. A crude derivative is taken on the filtered data, and any change from positive to negative on values above the noise floor is taken to indicate a peak.

The variations of the elevation angle (Figure 6(d)) and azimuth (Figure 6(e)) with time are instructive, since they highlight mode changes and their effects on the measured azimuths. The mode change marked in Figure 6(d) is from F1 to E to F2 modes. Figure 6(e) shows the associated changes of azimuth which are also clearly illustrated in Figure 6(f), which shows the number of occurrences of the angle of arrival in 1 degree bins in elevation and azimuth. In this particular case, the bearing increases with elevation angle, the higher-angle modes becoming increasingly off-great-circle.

The information contained in elevation-bearing plots such as Figure 6(f) can be used to determine values of azimuth which are more accurate than those normally obtainable with a network of bearing-only DF stations. In general, one would select out the data for the propagation mode considered most likely to have travelled closest to the great circle path - the E mode for this case, or the F1 mode for the Adelaide-Townsville circuit of Figure 7.

Figure 8 shows the SKYLOC data for a real, but unusual, case which nevertheless illustrates the high resolution and accuracy available under ideal conditions. Over 80% of the observations fell into either of two 1 degree elevation bins, with a similarly tight distribution in azimuth. The mean observed azimuth was 294.6 degrees, in agreement with the correct value. The unusual propagation conditions existing at the time have been described by McNamara [1], who suggests that propagation was via an M mode, with penetration of the F2 region almost occurring on one of the hops.

The bearing may be even more narrowly confined when propagation occurs via the E layer, as illustrated in Figure 9.

## 4 ACCURACY OF MEASURED ANGLES

As a pre-requisite to the attribution of errors in transmitter location to errors in the ionospheric models and the ray re-tracing procedures, it is necessary to establish the accuracy of the measured bearing and elevation angles passed to the modelling procedures. Except under some conditions discussed by McNamara [1], the accuracy may be estimated reliably when propagation occurs via a blanketing sporadic-E layer.

Figure 10 shows the observed angle distributions for a transmitter at Berri, which lies 187km from the SKYLOC site, on a bearing of 75.1 degrees. The elevation angle of 46.2 degrees corresponds to a ground range of 197km, for reflection from the observed blanketing Es layer at 105km. Data taken on the horizontal dipoles showed a second elevation peak at 66 degrees, corresponding to 2Es propagation to a range of 184km. In all, 6 sets of such data were collected, yielding an average range of 185km and a standard deviation of 10km. The standard deviation of 10km corresponds to a 5km change in the height of the Es layer, well within the accuracy for the manual scaling of the ionogram and height changes from one ionogram to the next.

The average bearing for the 6 sets of Berri data was 74.8 degrees, with a standard deviation of 1.6 degrees, covering the correct bearing of 75.1 degrees. The spread in bearing illustrated in Figure 10(b) is typical of that obtained for short Es circuits.

The Berri results thus confirm the generally high level of accuracy of the angles measured by SKYLOC. The high resolving power of the instrument is confirmed by the data for case 138 (Figure 8) for which over 80% of the measured elevation angles fell in either of two 1 degree bins, and the standard deviation of the bearing estimates was less than 0.4 degrees.

### 4.1 Difficulties At Low Elevation Angles

Some of the Sydney observations made during the 1986 trials of SKYLOC yielded very low elevation angles [1]. For example, for target 129, the observed elevation angle was 2.9 degrees. Such low angles must be regarded as unreliable, if only because the observing site can never be exactly flat and horizontal. Besides, the aperture projected at low angles is extremely small, and cannot provide the same accuracy as can be achieved with high-angle rays. Closer study of the data showed that the data-processing procedures used by SKYLOC cause the observations to be quantized for low elevation angles when the target lies almost along one arm of the array, making them even more uncertain.

This quantization is clearly evident in the data taken on the 5th March 1987, for a transmitter at Belconnen (Canberra; range 953km, bearing 96.3) on 6.4405 MHz (Figure 11) and 8.4675 MHz (not shown). The data cover approximately 9 hours of continuous collection, enhancing the visibility of the effect. The separations between the contours at a given azimuth for 8.4675 MHz were much smaller initially and decreased more rapidly with elevation angle than those for 6.4405 MHz, indicating that these contours are a function of frequency.

The contours arise because the processing performed within SKYLOC uses integral values of phase, on the assumption that the system is accurate to the nearest degree in phase. The use of integer arithmetic also offers advantages in speed of computation. The effect is particularly marked when the signal is arriving at a low elevation angle (< 10 degrees) and at an azimuth close to one of the arms of the array (in this case, near 90 degrees). With this geometry, an appreciable change in elevation angle occurs for each one degree in phase; hence integral phase values produce quantization of the elevation angle. Lower frequencies show much larger steps between the contour lines.

The measured azimuths are not greatly affected - the value obtained for target 129 was 88.0 degrees, less than a degree from the correct value of 87.2 degrees.

## 5 CONCLUSION

Techniques for the measurement of the angle of arrival of HF skywaves have progressed to the point where the accuracy of derived HF transmitter positions is controlled almost entirely by our limited ability to model the ionosphere along the circuit to the transmitter. The observations made by SKYLOC are of such high accuracy and resolution that they also provide a wealth of information concerning HF propagation, and make SKYLOC a powerful diagnostic tool for the study of the ionosphere for its own sake.

## 6 REFERENCES

1. McNamara, L.F. "Ionospheric modelling in support of single station location of long-range transmitters". Accepted for publication by J Atmos Terr Phys, 1988a.

2. McNamara, L.F. "Ionospheric limitations to the accuracy of SSL estimates of HF transmitter locations". These Proceedings, 1988b.

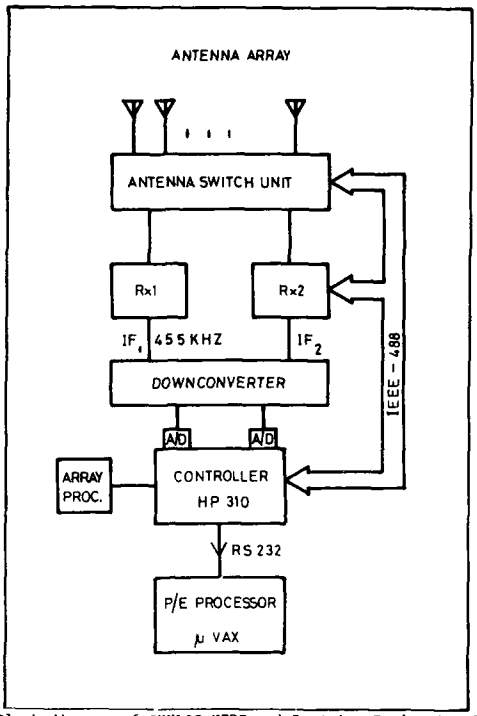



Fig 1. Block diagram of SKYLOC HFDF and Position Estimation System.

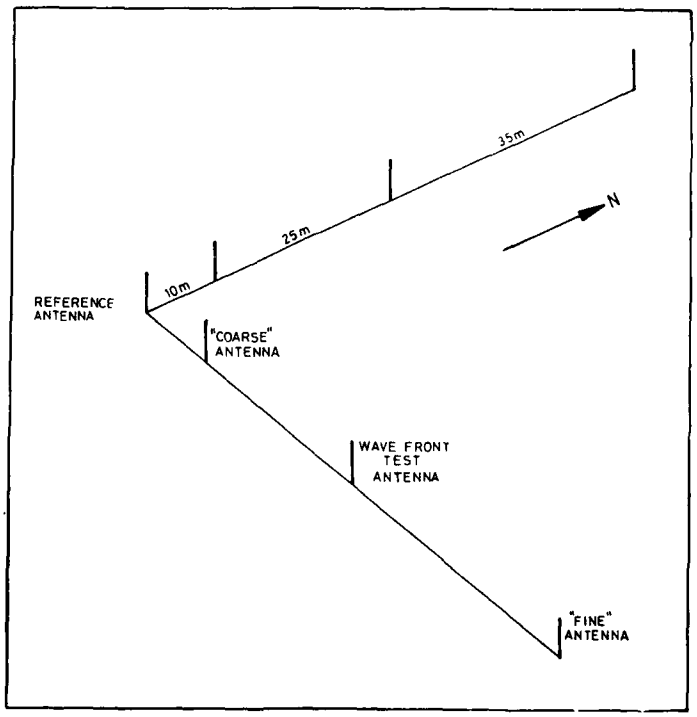



Fig 2. Typical antenna disposition used with SKYLOC for measuring the total angle of arrival of HF skywaves.

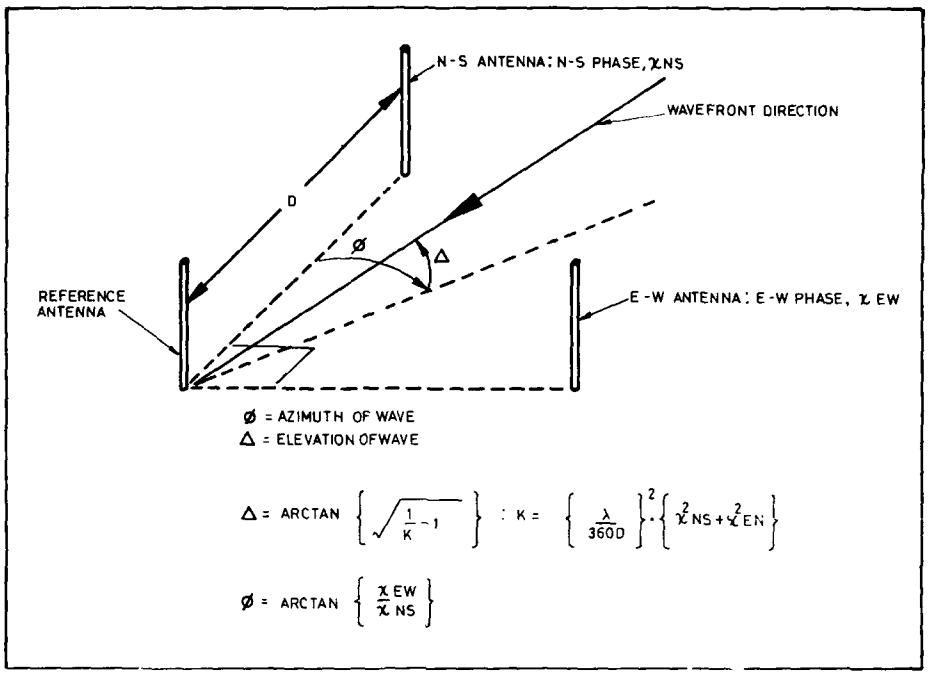



Fig 3. Procedures for determining the elevation angle and azimuth of a downcoming HF skywave.

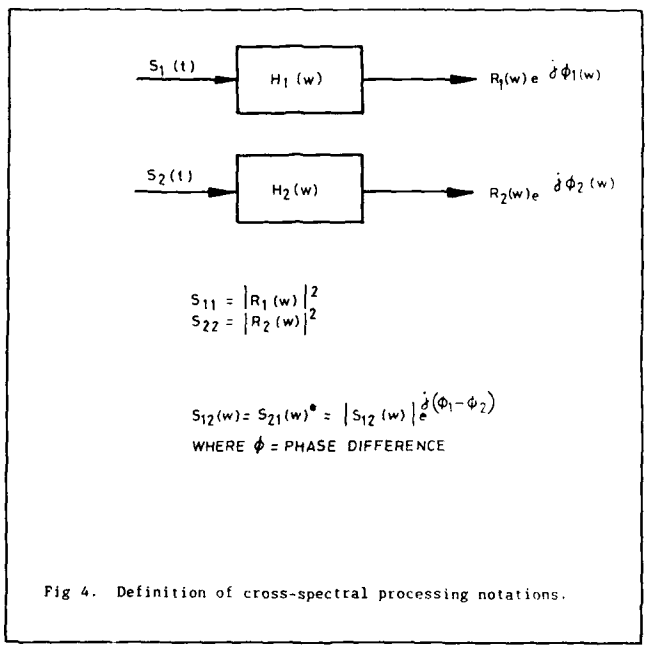



Fig 4. Definition of cross-spectral processing notations.

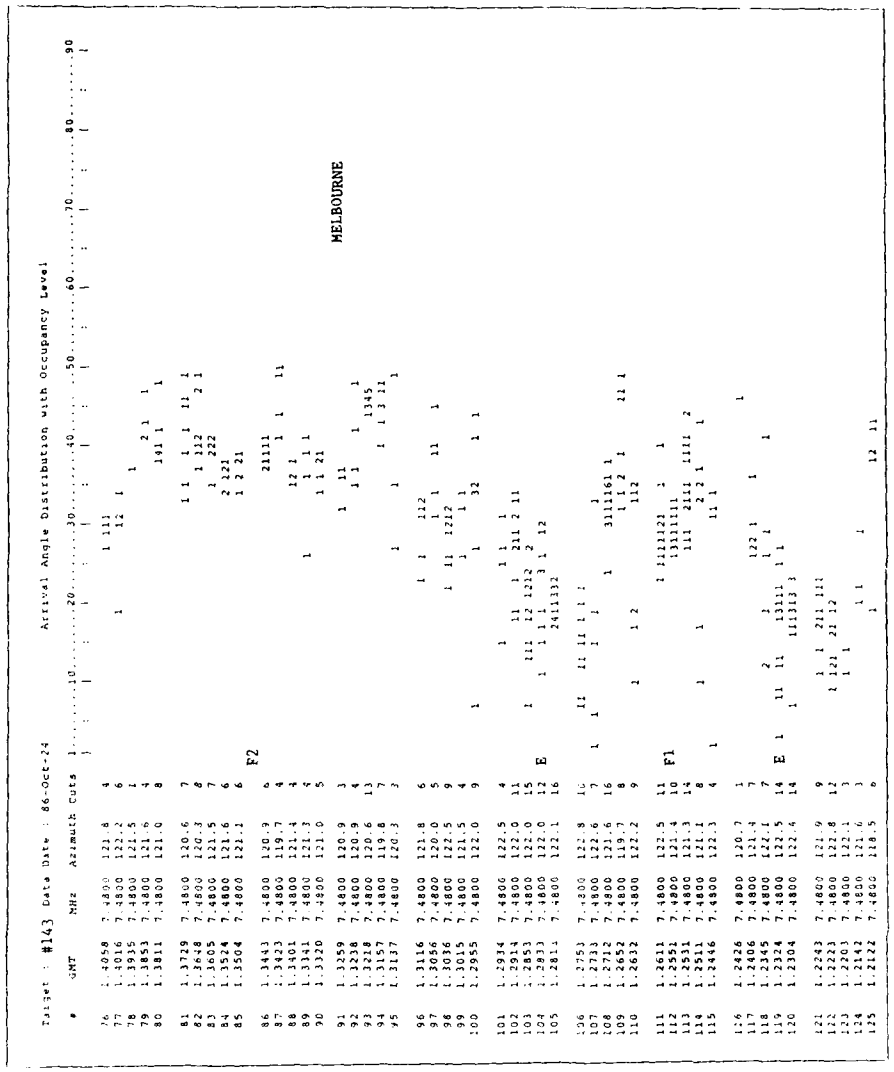



Fig 5. Numerical histograms of the elevation angles recorded for each 19-second group of data. The Greenwich mean time is given in hours, minutes and seconds. The total count for each group is nominally 19, but only those values passed by the various filters are counted. The azimuth is the mean value for the number of reliable "cuts". The mode identifications have been made using the one-dimensional quasi-parabolic model. This figure is not typical of those obtained during the SKYLOC trial - see text.

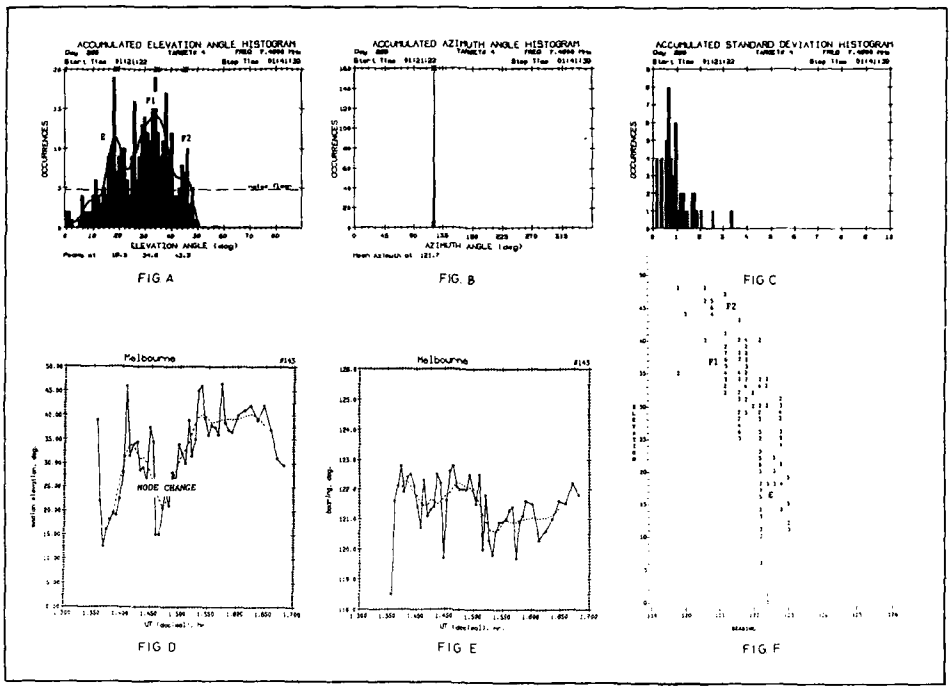



Fig 6. Accumulated (a) elevation angle, (b) azimuth, (c) standard deviation histograms; time series of (d) elevation angle and (e) azimuth; and (f) elevation versus bearing numerical histograms for Melbourne case 143.

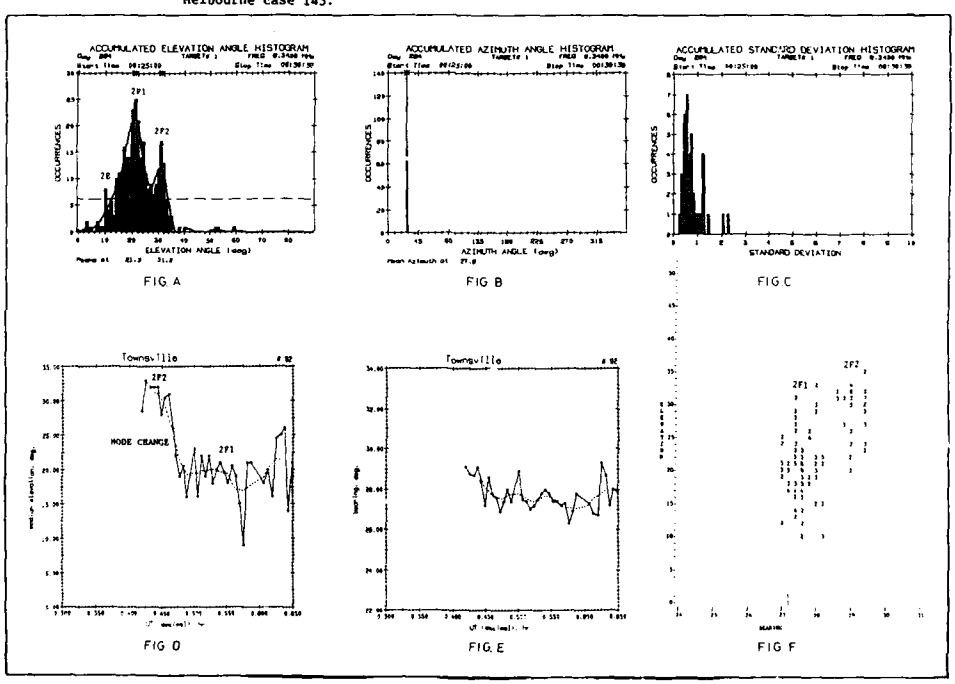



Fig 7. Accumulated (a) elevation angle, (b) azimuth, (c) standard deviation histograms; time series of (d) elevation angle and (e) azimuth; and (f) elevation versus bearing numerical histograms for Townsville case 92.

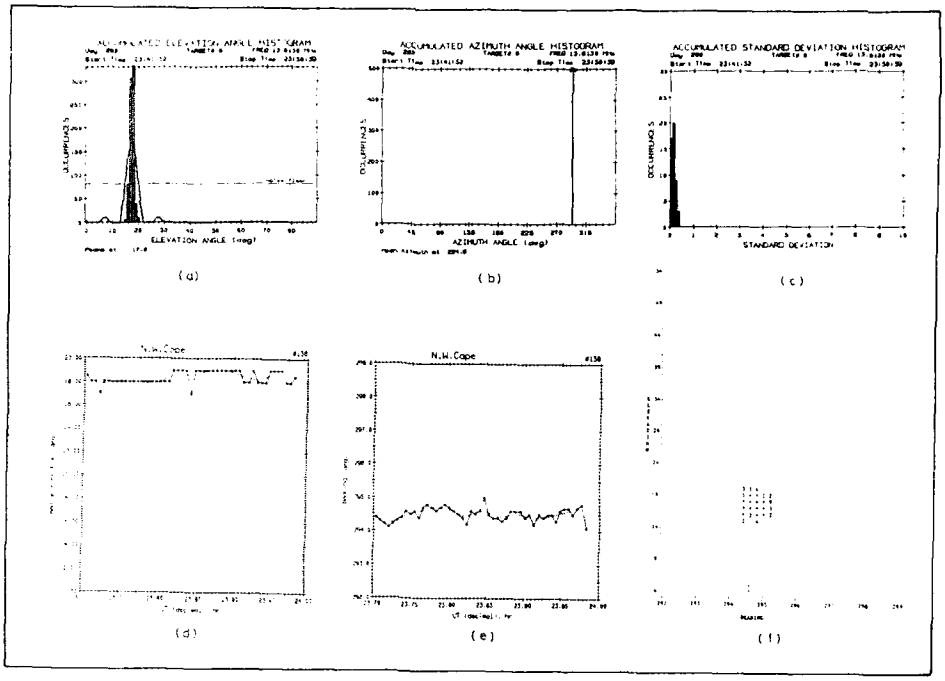



Fig 8. Accumulated (a) elevation angle, (b) azimuth, (c) standard deviation histograms; time series of (d) elevation angle and (e) azimuth; and (f) elevation versus bearing numerical histograms for Northwest Cape case 138.

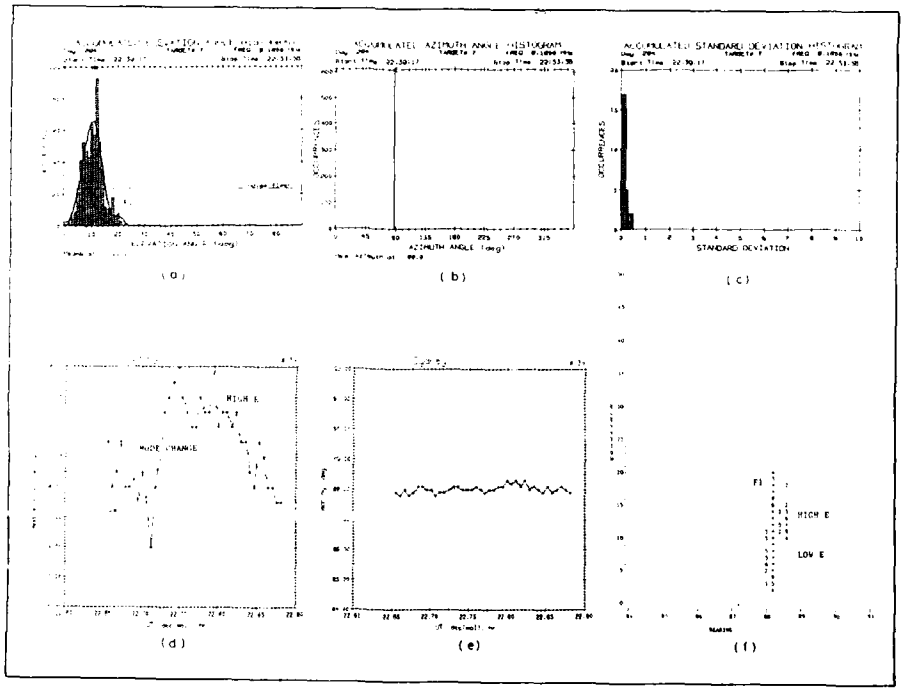



Fig 9. Accumulated (a) elevation angle, (b) azimuth, (c) standard deviation histograms; time series of (d) elevation angle and (e) azimuth; and (f) elevation versus bearing numerical histograms for Sydney case 39.

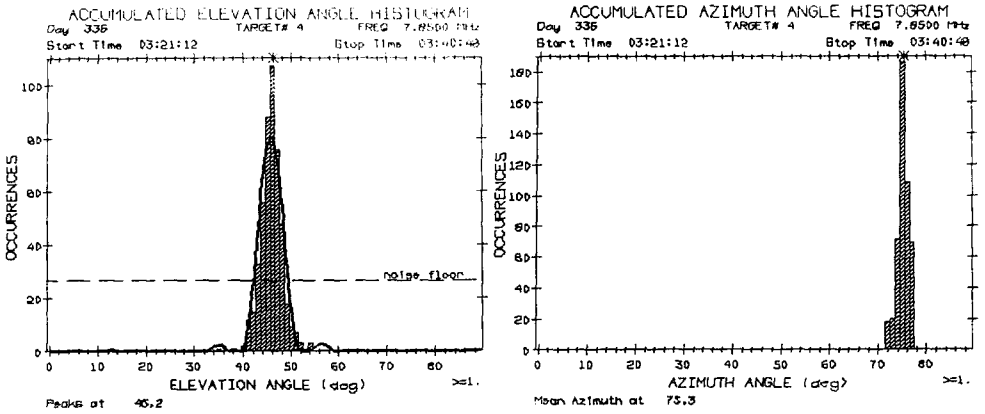



Fig 10. Accumulated histograms of elevation angles and azimuths for Es propagation from a transmitter 187km from the SKYLOC site, on a bearing of 75.1 degrees.

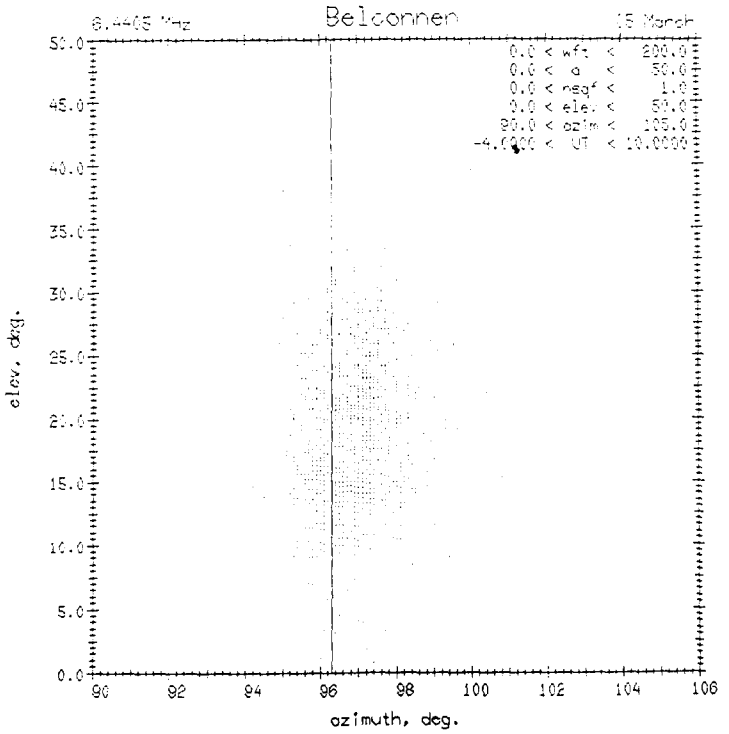



Fig 11. Elevation angle versus azimuth plot for data on a Canberra-Adelaide circuit accumulated for a period of about 9 hours. The striations at low elevation angle are caused by rounding off the observed phases to the nearest integer.

# IONOSPHERIC LIMITATIONS TO THE ACCURACY OF SSL ESTIMATES OF HF TRANSMITTER LOCATIONS

by

L.F.McNamara
Andrews Antennas
Innovation House
Technology Park 5095
Adelaide
Australia

**ABSTRACT**

Procedures for determining the location of an HF transmitter using observations from a single DF site have improved greatly in recent years. The ability to measure angles of arrival of the incoming skywave has now progressed to the point where the accuracy of the position estimation is set by the ionosphere and its variability, and by our imperfect knowledge of these variations, rather than by errors in the measured angles of arrival. After discussing the levels of accuracy currently obtainable, this paper discusses some of the limits set by our inability to specify the ionosphere accurately enough, and suggests methods by which some of these limitations can be overcome.

## 1 INTRODUCTION

Single Station Location (SSL) is one approach to the problem of the determination of the location of HF transmitters. The traditional method used to locate HF transmitters uses a network of at least two bearing measuring stations, with the estimate of position being obtained from the intersection of the two or more bearing estimates. An SSL station makes angle measurements in the elevation plane as well. The location of the transmitter is then specified as lying on the measured bearing, and at a range determined by tracing a ray at the observed elevation angle back towards the transmitter, using appropriate ionospheric models and raytracing procedures.

In general, the SSL technique can be regarded as complementary to the network approach, since each has its limitations. The SSL technique comes into its own when the networking technique fails because only one site can hear the transmitter, or when there is some confusion as to whether the sites are observing the same transmitter. The SSL technique, on the other hand, cannot always provide good range estimates because of inadequate knowledge of the ionosphere along the circuit, or because of multihop propagation.

There are three regimes in the elevation angle plane that require different approaches to the derivation of ground range. Above about 75 degrees, for short-range targets, it is necessary to take account of the detailed motion of the ionosphere with particular regard to the effects of travelling ionospheric disturbances (TIDs). Between about 75 degrees and 50 degrees, for medium-range targets, most of the effects of TIDs will have been integrated out as the ray passes through the ionosphere, and the local ionogram is usually valid for the calculation of transmitter range. Below about 50 degrees, for long-range targets, the local ionogram is only indicative of what is happening to the ionosphere along the circuit, and recourse must be made to synoptic models of the ionosphere.

The present paper is mostly concerned with long-range targets, since it is here that extensive knowledge of large parts of the ionosphere is required, and briefly concerned with medium-range targets. For short-range targets, mathematical techniques based on an understanding of the small-scale effects of TIDs are required, but these will not be discussed here.

The paper discusses some of the limitations of the SSL technique, with particular reference to the requirement for a reliable real-time model of the ionosphere along the circuit and for correct interpretation of the propagation conditions along the circuit. Discussions are couched in terms of the results obtained with the Andrew SSL system SKYLOC [Trade Mark] during a series of trials with uncooperative transmitters.

## 2 RAY RE-TRACING

The computational procedure in which rays are traced from the DF site back towards the transmitter through a model of the ionosphere, at the measured bearing and elevation angles, is known as ray re-tracing.

The ray re-tracing technique relies on having a sufficiently accurate model of the ionosphere along the whole path between the transmitter and the DF site. This model would specify the electron density profile N(h) at all points along the path, even when the path is not a great-circle path. In the ideal case, an ionogram is available for every point along the path, and this ionogram is readily interpreted and analysed to deduce a reliable N(h) profile. At the other extreme, no ionosonde information is available and recourse must be made to maps of ionospheric characteristics based on historical data, and to some reasonable N(h) profile "anchored" by these characteristics. The latter approach will fail completely if the ionosphere is in any way abnormal.

Somewhere in between these limits is the case most often encountered in practice - some ionograms are available, but not exactly where they are required. It is then necessary to use information from these ionograms to modify the historical maps so that they are more applicable to the current situation.

The process of setting up the appropriate model of the ionosphere can be divided into three stages:-

1. Obtaining suitable world-wide synoptic models (either formulas or maps) of the required ionospheric characteristics, with some method of updating to match the ionosphere observed at the time

2. Making observations of the ionosphere where possible, to determine how to update the synoptic models

3. Erecting an appropriate electron density profile which matches the updated ionospheric characteristics.

Generally speaking, it is necessary to define three characteristics for the E, F1 and F2 layers - the critical frequency fc, the height at which fc occurs, hm, and a "thickness" parameter ym. All parameters cannot be defined with equal accuracy. The accuracy decreases from fc to hm to ym, and from the E to F2 to F1 layers.

One way of using ionosonde observations at one location to update synoptic-map values at another location is to use an effective ionospheric index. A different index should be used for the F2 layer from that used for the E and F1 layers, since the causative physical mechanisms are different and independent. Attempts have been made to relate departures of ionospheric characteristics from the median at one location to those at another location, but the published correlations seem to have been overly optimistic [1,2].

The problem of ray re-tracing may be approached at many levels of complexity, depending on how well the ionosphere can be modelled and how much computational power is required. There will sometimes be a trade-off between how quickly and how accurately the results are required, and there is never any point in performing slow but accurate raytracing through a model of the ionosphere which is just an educated guess.

In order of increasing complexity, some of the methods which may be adopted are:-

1. One-dimensional ionosphere; no magnetic field; analytic raytracing. This method is called the one-dimensional quasi-parabolic or [1]QP method.

2. One-dimensional ionosphere; with magnetic field; numerical raytracing.

3. Two-dimensional ionosphere; no magnetic field; analytic raytracing. This method is called the two-dimensional quasi-parabolic or [2]QP method.

4. Two-dimensional ionosphere; with magnetic field; numerical raytracing.

5. Three-dimensional ionosphere; with magnetic field; numerical raytracing.

The results presented in this paper were obtained using the methods 1, 3 and 5, covering the range from least to most complicated.

## 3 RESULTS OF SYSTEM TRIALS

Before proceeding to a discussion of some of the limitations of the SSL method of position estimation, it is appropriate to discuss just how good the method can be in practice.

A trial of the SKYLOC system was run for a week in October 1986, and then for two days in the following December, with unknown transmitters located at more than twenty different sites around Australia. The trials were coordinated by the Australian Department of Defence on behalf of that Department and various US Government Agencies. Approximately 250 targets were provided to the SKYLOC operator, each being on the air for a nominal 15 minutes and identified only by the frequency [illegible] modulation type. The operating time was 2000UT to 1200UT, covering both the pre-sunrise and post-sunset periods. The trials were conducted very near solar minimum, and there were no solar-related ionsopheric disturbances.

SKYLOC was located at a site within the boundaries of the Defence Science and Technology Organization, just north of Adelaide. The site suffered from at least a 10dB signal-to-noise disadvantage over the quiet receiving site used by the trials coordinator 10km away, causing some difficulties with target acquisition and identification. Unfortunately, it was not possible to provide the transmitters with a unique identifier. In general, the spectral processing in SKYLOC allowed it to work successfully down to at least -5dB S/N, below which the lack of a target identifier made it impossible to be sure that the required transmitter was being observed.

The lack of an identifier caused some confusion even in the case of high S/N ratios. For example, target 117 had an S/N ratio of 17dB, and was supposed to be Sydney. However, the transmitter observed was determined to be at Macquarie Island (bearing 150.5, range 2690km). Generally speaking, the ionosphere on the circuits considered during the five-day October trial was a typical mid-latitude ionosphere, with normal diurnal variations, the occasional TID and seasonal sporadic E (Es). There were no features which would have made position finding any harder or any easier than at any other mid-latitude site.

Blanketing Es persisted throughout most of the two-day December trial, making range estimation an easy task while at the same time permitting validation of the elevation angles measured by SKYLOC [3].

McNamara [4] has analysed the October trials data for the long-range targets, concluding that it is possible to obtain median absolute miss distances of around 7% for both E and F propagation modes, using synoptic models of the ionosphere updated by effective ionospheric indices derived from a single ionosonde. However this analysis presumed a knowledge of the order of the propagation mode, with many of the longer circuits supporting 2-hop modes. RMS bearing errors were 0.8 degrees on average.

Tables 1 and 2 show the errors in the calculated ranges for modes identified as normal E modes (Table 1) and F2 modes (Table 2). The E-mode results are for the multi-segmented quasi-parabolic profile used by McNamara [4], with analytic raytracing ignoring the earth's magnetic field. Elevation angles flagged as negative in Table 1 were associated with secondary peaks in the elevation angle histograms, the E mode not being the dominant mode.

Three sets of errors are given for the F2 cases listed in Table 2. The first method, [1]QP, uses the quasi-parabolic profile, with iteration to determine the location of the reflection point and hence the values of the ionospheric characteristics; the second method, [2]QP, allows the quasi-parabolic profile to vary along the circuit; and the third method uses a full three-dimensional model of the ionosphere, with a Dudeney [5] N(h) profile and raytracing with field [6].

It can be seen that there is little to choose between the three methods, the size of the error being associated with the target rather than with the method of position estimation. This indicates that accuracy limits on position estimation for F2 modes are currently being set by our inability to obtain a sufficiently accurate model of the ionosphere along the circuit, and not by our inability to simulate the propagation through that model.

## 4 BASIC LIMITATIONS OF THE APPROACH

Two basic limitations will be discussed here, with particular reference to long-range transmitters. The first is a limitation of the SSL technique itself, while the second is a limitation of the method used to analyse the observations of the incoming wave front.

### 4.1 Determining The Order Of The Propagation Mode

There is a conventional wisdom in some parts of the DF community that the order of a propagation mode may be determined from the spread in the observed bearing of the signals. This belief seems to have arisen from the observation that the bearing deviations are greater for multi-hop modes than for single-hop modes [7]. This is what is expected logically and is supported by observations made during the October trials of SKYLOC. For example, the observations of target 204 (Brisbane) showed a 2F2 mode on a bearing at 63.8 degrees with a standard deviation of 0.6 degrees, while a 3F2 mode on a bearing of 62.9 degrees had a standard deviation of 1.8 degrees, three times as great as that for the 2F2 mode.

However, the reverse is not true - a large spread in bearing does not uniquely brand a mode as multi-hop. In fact, the smallest spread in bearing for all targets observed during the two trials was for a 2-hop F2 mode (Target 138, North West Cape). As discussed later, the spread in bearing is usually set by the presence of several propagation modes, each with its own bearing distribution.

The lack of any information regarding the order of the propagation mode is a basic limitation to the SSL technique. However, the order can often be deduced by using ancillary information such as land/sea interfaces, population centres, country borders and so on.

### 4.2 Failure To Obtain QUMP

Being based on a wave-front test procedure rather than a wave-front analysis procedure, SKYLOC relies on the ability to identify the existence of periods for which all but one propagation mode fade away, leaving effectively just one propagation mode. This is the condition of quasi uni-modal propagation or QUMP [8].

This condition was easily met during the trials on those circuits which supported only a 2F2 mode - Darwin-Adelaide and North West Cape-Adelaide. It was also met to a sufficient degree for most other circuits, depending on the number of modes which were supported by the ionosphere and the radiation pattern of the transmitting antennas. QUMP is more readily obtained when there is only one propagation mode possible, or when the transmitting antennas favour one mode.

Figure 1 shows the accumulated elevation angle histogram for a case in which a 20-minute observation period was insufficient to obtain a useful histogram. The problem was not due to an inadequate S/N ratio, since the observed value of +11dB was more than 20dB above the level at which SKYLOC can work. The intended transmitter was located at Hobart at a range of 1157km from Adelaide, and on a bearing of 142.5 degrees. The observed bearing of

141.8 degrees is in good agreement with the correct bearing. Propagation simulation indicated that the operating frequency of 5.855MHz would have been supported by a 1E mode (6 degrees), a 2E mode (18 degrees), a 2F1 mode (33 degrees) and a 2F2 mode (46 degrees). With this number of propagation modes possible, it is not surprising that long periods of QUMP were rare. In cases such as this, there is clearly room for wavefront analysis procedures. The curve through the histogram was defined by an automatic peak-picking procedure [3].

## 5 LIMITATIONS TO RANGE ESTIMATION FOR NORMAL MODES

While the techniques developed to date provide acceptably accurate range estimates in general, there will be some conditions under which accurate range estimates cannot be obtained:-

1. Estimates are unreliable if associated with a cusp, especially for a high ray
2. The base of the E layer is poorly modelled
3. The N(h) profile in the E-F2 and F1 regions is poorly modelled
4. Synoptic models do not correctly model large-scale tilts in the ionosphere
5. There is no way to determine if propagation on a long circuit was by the ordinary or extraordinary ray, and magnetic fields are sometimes important
6. There is no "correct" way to alter a synoptic model when that model predicts that the observed propagation is not possible.

Figure 2 shows a range versus elevation angle plot for the simulation of an 8.47MHz signal propagating from Canberra to Adelaide at around 1300LT. The plot is based on a simple 3-layer model of the ionosphere, with analytic raytracing and ignoring the earth's magnetic field. The ground range to the actual transmitter was 950km, and there were peaks in the observed elevation angle histogram at 9 degrees and 29 degrees. The model yielded quite accurate and fairly independent range estimates for the 1E and 1F2 propagation modes.

The figure serves to illustrate some of the propagation phenomena which must be taken into account in SSL position estimation. Each of the layers has a corresponding skip distance, high and low rays, and penetration above the MUF. A certain amount of skip-distance focussing is indicated by the figure for the E layer (17 degrees, 680km); F1 layer (22 degrees, 940km) and F2 layer (35 degrees, 800km). The effect would be greatest for the F2 layer, which has the smallest variation of ground range versus elevation angle. The high rays for the E and F2 layers have a rapid variation of ground range with elevation angle, and would probably be substantially defocussed in practice.

QUMP would be easiest to achieve for ranges less than 800km (only one mode, the E mode, is supported), and then less than 940km (two modes, E and F2 modes, are supported). Circuits with a length of 1000km would support at least 1E, 1F1 and 1F2 modes, with probably only the low E and low F2 modes. Both the low and high 1F1 modes may exist, but the peaks in the elevation angle histogram would probably not be resolved. When this number of modes is supported, periods of QUMP tend to be fairly limited in duration.

### 5.1 Effects Of Cusps

There will clearly be some difficulty if an observed elevation angle peak corresponds to a cusp in the range-elevation plot, since the ground range is then essentially undetermined. There is also the difficulty that the simulated cusp could be quite different from the cusp for the actual ionosphere, especially on the steeper high-angle side. In practice, defocussing and enhanced deviative absorption would make it unlikely for any observed signal to correspond to the steepest parts of the cusps.

### 5.2 Base Of The E Layer

A study of Table 1 shows that the largest errors for E-mode propagation occurred for the Sydney-Adelaide circuit and for low elevation angles. Although the values of the measured elevation angles for these cases are suspect [3], they draw attention to the very low apogee heights reached by the rays on that circuit, especially around dawn. At these heights, the present parabolic E-layer model starts very abruptly from the base of the layer and a small change in the base height can have a large effect on the calculated range.

Ideally, the base of the model E layer would have a more complicated structure, to match the observed D and lower E region N(h) profile, but our present knowledge of these regions is inadequate for this task.

### 5.3 The F1 Layer

If the ionospheric model used for the simulation in Figure 2 had been a perfect match to the real ionosphere, the circuit should have been observed to support propagation by the E, F1 and F2 layers. The fact that no elevation peak was observed near 22 degrees shows

3. It is not possible to allow for simple Es or Es/F combination modes reliably.

4. For low frequencies and high obliquity factors, it is possible to get reflection from the residual Es layer at 95km.

### 6.1 Es Mode Identification

If the only ionosonde available is one at the DF site, any observations with regard to Es can be assumed to be applicable only for target ranges of up to about 200km, since individual Es patches are usually only a few hundred km in extent. The duration of the Es echoes at the DF site would give some indication of how localized the Es probably is, since any long-lived ionospheric phenomenon can usually be assumed to cover a large area, but this effect has not been quantified to date.

Even when Es does cover the DF site and the circuit reflection point, the detailed characteristics of the Es may be quite different. In particular, the Es may be non-blanketing at the DF site, and blanketing at the circuit midpoint, or vice versa. In the former case, it would be appropriate to quote two target ranges - one assuming Es propagation and one assuming a normal reflection mode.

### 6.2 Non-blanketing Es

There were two cases recorded during the October trial that propagation on the Canberra-Adelaide circuit relied on an Es layer. These were cases 108 and 123, and in both cases the assumption of propagation by normal modes would have placed the transmitter well off the east coast of Australia instead of at Canberra. The Adelaide ionogram in fact showed well developed Es on both occasions. For case 108 (Figure 4), the observed elevation angle was 11.8 degrees, corresponding to a range of 940km for an Es layer at the observed height of 120km, in excellent agreement with the correct range.

For case 123 (Figure 5) on the other hand, the observed peak in the elevation angle histogram of 8.1 degrees requires an Es layer at 90km to yield the correct range. The Es observed at Adelaide lay at a height of 120km, with no lower Es trace with a plasma frequency exceeding the minimum observed frequency of 1.8 MHz. J D Whitehead [10] suggests that the propagation in this case may have involved reflection from the tilted edge of an Es cloud. This interpretation is consistent with the much wider spread in the measured elevation angles for case 123 than for case 108, and is also consistent with the partially blanketing nature of the Es observed at Adelaide. However it seems unwise to push the interpretation too far, because of the small scale size of Es clouds relative to the length of the circuit. The spread in bearing is essentially the same for the two cases.

### 6.3 Es Combination Modes

The presence, or probable presence, of Es near the DF site, near the target, or near one of the ground reflection points makes it impossible to identify the propagation mode in some cases. For example, instead of a 1F2 mode, the mode could be Es-F2, F2-Es and so on. This ambiguity is another facet of the inability of an SSL system to identify a multi-hopped propagation mode. In the case of a 2-hop mode, it is possible for the ground near the midpoint to be shielded by an Es layer, resulting in an M reflection. This appears to have happened for case 138, for a transmitter at North-West Cape. The estimated ranges were about 1800km for a 1F2 mode and 3300km for a 2F2 mode, depending on the model, straddling the correct value of 2783km. However when the ground reflection for the 2-hop mode is replaced by an M reflection from the top of an Es layer (placed at hmE for computational convenience), the range for the one-dimensional quasi-parabolic model becomes 2939km, an error of only 6%. The error rises to 17% if the propagation mode is taken as EsF2.

The interpretation of this case in terms of an M reflection is supported by the results from Perth case 134 an hour earlier, in which the calculated 1F2 and 2F2 ranges were about 1400 and 2800km respectively. Taking the mode to be an EsF2 mode, with the Es over Western Australia, yields a range of 1903km for the quasi-parabolic model, an underestimate of only 11%.

The results presented in Table 2 exclude all cases for which the models yield essentially the same range, but this range falls short of the transmitter by the range of a 1-hop Es mode. There seems to be no characteristic of the observations that would identify the propagation mode as involving Es.

## 7 LIMITATIONS TO BEARING ACCURACY

The measured bearing for the long-range targets considered during the October and December 1986 trials had a systematic error of 0.6 degrees, and an RMS error of 0.8 degrees. The systematic error was due partly to injudicious handling of the data, and probably partly due to site errors. The former has since been removed, but the latter has not been investigated because there was no permanent DF site.

that Canberra was outside the skip zone for propagation via the F1 layer, and that the F1 layer model was incorrect.

Our generally poor level of knowledge of the F1 region not only results in inaccurate range estimates for F1 modes, but also often prohibits the reliable identification of an elevation peak as corresponding to an F1 mode. The difficulty lies not with foF1, which is well modelled, but by inadequate models for hmF1 and ymF1. The models currently used appear to have been based on only a limited set of observations, and are not always representative of the actually ionosphere. These limitations are compounded by the very real difficulty that the F1 layer often does not exist as a separate layer.

### 5.4 Polarization Ambiguity

Whereas the crossed loop or horizontal dipole antennas used for short-range targets allow the selection of either the ordinary (O) or extraordinary (X) ray, the vertical monopoles used for long-range targets do not. This means that the elevation angle distribution for the 1F2 mode, for example, will normally show just one peak, which in reality is an unresolved double peak corresponding to the two modes of polarization. Simulations of this effect have been performed by Saout and Bertel [9].

Some evidence for the presence of the two polarization modes can be obtained by studying how the histograms build up, rather than just the histogram accumulated over the whole observing period. Figure 3 shows evidence of a 1F2 X mode at around 36 degrees elevation and in O mode at around 41 degrees, during the first few minutes of observation of a transmitter in Melbourne. At the end of the total 17 minute observing period, the histogram was single-valued, with a peak at around 40 degrees. This peak corresponds most closely to the O mode, but in general there would be no way of making this identification.

The calculated ranges assuming a Dudeney N(h) profile are 505km for the X ray, and 515km for the O ray. The no-field range corresponding to the 40.5 degrees peak in the accumulated histogram is 500km, in good agreement with the other calculated ranges. The agreement would not have been so good if the X ray had been the dominant contributor to the final elevation histogram, since the X-mode range for an elevation angle of 40.5 degrees is 435km. The fact that all ranges fall short of the correct range of 630km shows that either the ionospheric characteristics, or the Dudeney profile shape for the F2 layer, was not a good fit to the real ionosphere in this case. The calculated E-mode range was 646km.

### 5.5 Use Of Check Targets

The updating of a synoptic model of the ionosphere along a circuit using the gross properties of the ionosphere observed at any available ionosonde site relies on the validity of the assumption that the available ionospheric observations are representative of the ionosphere along the circuit. Clearly this assumption has its limitations, and where possible it is preferable to sample the ionosphere along the circuit itself.

This can best be done using an oblique ionosonde, but it is highly unlikely that one will be operating on a circuit sufficiently close to the required circuit to offer much improvement to the ionospheric model. It is more likely that a known fixed-frequency transmitter exists in the general region of the target, and observations on this known transmitter (or check target) can be used to modify the model so that it yields the correct position (range and bearing) for the known check target.

There appears to have been no detailed study made of this problem, but the main requirement for a known transmitter to be a useful check target is that the check target and unknown target should be as close as possible in location and operating frequency. It is a matter of experiment to decide just how far these restrictions can be relaxed, and under what conditions, but it is essential that reflection should at least occur in the same layer for both transmitters.

It is quite possible that no reasonable amount of modification of the synoptic model will enable it to locate the check target accurately, because the model does not adequately represent the horizontal gradients present in the ionosphere at the time of the observations. In such cases, it may be preferable to revert back to the simpler model of tilted plane mirrors, and work with fewer degrees of freedom.

## 6 DIFFICULTIES CAUSED BY SPORADIC E

While under some conditions the presence of Es can make the procedure of range estimation quite accurate, it does cause some problems:-

1. There is no reliable method for determining whether reflection on even a one-hop circuit was via Es.

2. Non-blanketing (patchy) Es can give rise to tilts at the edges of patches, and these give rise to incorrect range estimates.

Examination of the observed histograms of bearing show that the histogram often does not cover the correct bearing, the discrepancy being greater than the resolution of the system. Discrepancies between true and measured bearings can be attributed to either large-scale tilts, as at sunrise or sunset, or to small-scale tilts due to TIDs. The latter would tend to average out as the period of observation increased beyond the typical TID periodicity of 10 to 20 minutes. The observed discrepancies are accumulated over about a 20 minute period, and their size depends on whether or not a TID was present on the circuit, the size and direction of the TID, and whether or not the observing period covered a whole cycle of the TID.

Figure 6 shows by way of example the variation of bearing with time for target 47 (Perth), and the variation of elevation angle with bearing. Asterisks denote counts greater than 9. Figure 6(a) shows evidence of a large TID with a period of about 15 minutes which caused a bearing swing of about 1 degree, while Figure 6(b) shows that the TID caused both the bearing and elevation angles to change, and not in a systematic fashion. The weighted mean bearing of 273.6 degrees is 1.3 degrees greater than the correct bearing, possibly due to a morning tilt in the F2 layer since the observing period covered one full cycle of the TID.

### 7.1 Bearing Corrections

It might be expected that some of the discrepancies could be removed by allowing the re-traced rays to move off the great circle in response to transverse gradients in the synoptic ionospheric model. However this expectation is not met. Figure 7 shows the calculated bearing errors, obtained by raytracing through a three-dimensional model of the ionosphere with magnetic field, versus the observed bearing error, for F2 propagation modes. There is obviously no correlation between the two sets of discrepancies, indicating that synoptic models of the ionosphere do not model the horizontal gradients in the ionosphere accurately enough to permit the determination of bearing corrections by ray re-tracing.

### 7.2 The Ross Curve

The Ross curve [11] shows the variance in bearing as a function of target range, as derived from a large amount of good quality data mostly collected during World War II. The variance represents the ionospheric effect on the bearing error due to disturbances such as TIDs. There is a general decrease of the variance with range out to about 1000km, followed by an increase attributed to the presence of several propagation modes on the circuit.

The bearing errors obtained during the October trial were found to lie below the Ross curve in general, as illustrated in Figure 8. The circuit locations lying closest to the Ross curve, Melbourne, Brisbane and Irirangi, usually supported at least two and often three propagation modes, each with its own mean bearing. The Darwin and North-West Cape propagation modes were almost always 2F2, showing that the spread in the bearing is not controlled by the order of the propagation mode.

The propagation mode for the Sydney circuit was usually a 1E mode, with none of the bearing spread characteristic of F-layer propagation. Consequently the data point for Sydney lies well below the Ross curve.

## 8 MEDIUM-RANGE TRANSMITTERS

In many ways, determining the locations of transmitters lying at medium ranges (that is, about 200 to 500km) is the simplest of all cases, because a local ionosonde can be used to observe and quantify the ionosphere, while the effects of small-scale TIDs are not yet overwhelming. Ideally, (a) the local ionogram would apply exactly to the reflection point for the unknown transmitter, (b) the ionogram would be easily analysed to obtain the electron density profile, N(h), and (c) there would be no large-scale tilts of the ionosphere. The location of the transmitter could then be determined by raytracing through the derived N(h) profile, including the effects of the earth's magnetic field.

The most likely limitation to this overall procedure is an inability to interpret the local ionogram (preferably done automatically) and thence derive the N(h) profile. There will also be the occasional large TID which will necessitate the use of a check target to enable an accurate position estimate to be determined. For example, Figure 9 shows elevation and bearing versus time plots for a transmitter at Naracoorte (range 313km, bearing 142.6), which clearly indicate the passage of a large TID directly across the circuit. The effects of large-scale tilts could also be taken into account by the use of check targets.

Traditionally, the range estimation for a medium-range transmitter would be performed using the so-called "Classical SSL" technique. This technique is based on three fundamental laws of HF radio propagation - the Secant Law, Breit and Tuve's Theorem, and Martyn's Equivalent Path Theorem. The application proceeds as follows:-

1. The incoming signal at the frequency f is observed to have an elevation angle X.

2. The secant law is used to calculate the equivalent vertical frequency, fv.

3. The virtual height h'(fv) is read from a local vertical incidence ionogram, at the frequency fv.

4. The signal is assumed to have travelled a triangular path, with an elevation angle of X, and with reflection occurring at an altitude equal to h'(fv).

5. The range to the transmitter, TR, is then simply 2h'/tan X.

The usefulness and validity of this technique is limited by its assumption of no magnetic field, which introduces an inconsistency when h'(fv) is scaled from a real ionogram which clearly shows the effects of the earth's field.

The median absolute miss distance for medium-range targets observed during the October 1986 trials of SKYLOC was 9% for both the classical SSL technique and the one-dimensional quasi-parabolic approach, somewhat higher than for the long-range targets. The bearing errors were also somewhat larger than for the long-range targets.

## 9 CONCLUSION

Many of the limitations of the SSL technique can doubtless be overcome by more careful consideration of each of the potential problems raised here. Given that the accuracy and resolution of the equipment used to measure the total angle of arrival is at least one step ahead of our ability to model the ionosphere, it is obvious that the next move must be to improve that ability. This can only be done by further improvements in our knowledge of the physics of the ionosphere, and by real-time monitoring of the ionosphere on a large scale.

## 10 REFERENCES


1. Rush, C.M. "An ionospheric observation network for use in short-term propagation predictions." Telecomm J, Vol 33, 544, 1976.

2. McNamara, L.F. and Wilkinson, P.J. "A cautionary note on the use of F2-region correlation coefficients for short-term forecasting purposes." Solar-Terrestrial Proceedings. Proceedings of a workshop at Meudon, France, June 18-22, 1984. National Oceanic and Atmospheric Administration, Boulder, Co., and Air Force Geophysics Lab, Bedford, MA, 1986.

3. Jeffrey, Z.R., Middleton, P.T. and Winkler, C. "Accurate measurements of the total angle of arrival of HF skywaves." These Proceedings, 1988.

4. McNamara, L.F. "Ionospheric modelling in support of single station location of long-range transmitters." Accepted for publication by J Atmos Terr Phys, 1988.

5. Dudeney, J.R. "An improved model of the variation of electron concentration with height in the ionosphere." J Atmos Terr Phys, Vol 40, pp 195-203, 1978.

6. Jones, R.M. and Stephenson, J.J. "A versatile three-dimensional ray-tracing computer program for radio waves in the ionosphere." US Dept of Commerce Office of Telecommunications Report OT 75-76, 1975.

7. Barfield, R.H. and Ross, W. "The measurement of lateral deviation of radio waves by means of a spaced-loop direction-finder." J Inst Elect Eng, Vol 38, p 98, 1938.

8. Gething, P.J.D. "Radio Direction Finding". Peter Peregrinus Ltd, London, 1978.

9. Le Saout, J.Y. and Bertel, L. "SKYLOC system simulation". Private communication, 1988.

10. Whitehead, J.D. Private communication, 1988.

11. Ross, W. "The estimation of the probable accuracy of high-frequency direction-finding bearings". J Inst Elect Eng, Vol 94, p 722, 1947.

| CASE | TX | ELEV | APOGEE | RANGE | % ERROR |
|---|---|---|---|---|---|
| 3 | CANBERRA | 10.0 | 92.0 | 883 | -7 |
| 45 | (953 km) | 25.8 | 102.1 | 943. | -7 |
| 80 | | -8.0 | 104.2 | 1154 | 21 |
| 81 | | -11.0 | 96.5 | 898 | -6 |
| 84 | | 13.2 | 100.4 | 850 | -11 |
| 97 | | 14.4 | 102.2 | 828 | -13 |
| 111 | | 9.0 | 94.3 | 995 | 4 |
| 131 | | -10.0 | 97.4 | 979 | 3 |
| 141 | | -8.0 | 95.1 | 1090 | 14 |
| 35 | MELBOURNE | 16.0 | 95.6 | 646 | 2 |
| 48 | (633 km) | 16.0 | 96.0 | 651 | 3 |
| 53 | | 20.2 | 100.2 | 576 | -9 |
| 86 | | -13.0 | 92.4 | 725 | 15 |
| 102 | | -19.0 | 99.0 | 593 | -6 |
| 135 | | 17.6 | 102.0 | 681 | 8 |
| 143 | | 19.0 | 98.4 | 586 | -7 |
| 6 | SYDNEY | 4.5 | 88.7 | 1361 | 21 |
| 39 | (1124 km) | 7.8 | 94.0 | 1106 | -2 |
| 49 | | 6.8 | 88.6 | 1106 | -2 |
| 62 | | -8.0 | 95.3 | 1102 | -2 |
| 83 | | 3.9 | 94.6 | 1552 | 38 |
| 94 | | 6.4 | 88.6 | 1145 | 2 |
| 105 | | 7.8 | 99.5 | 1241 | 10 |
| 129 | | 2.9 | 96.4 | 1742 | 55 |
| 140 | | 6.0 | 94.3 | 1311 | 17 |
| 148 | | 7.5 | 89.0 | 1048 | -7 |
| 155 | | 6.7 | 100.3 | 1364 | 21 |
| 54 | PERTH | -7.0 | 94.6 | 2340 | 10 |
| 60 | (2127 km) | 11.4 | 101.6 | 1952 | -8 |
| 87 | | -10.0 | 96.8 | 1935 | -9 |
| 52 | BRISBANE | -14.0 | 97.7 | 1520 | -4 |
| 66 | (1579 km) | 6.0 | 107.6 | 1429 | -9 |
| 89 | | -4.0 | 92.9 | 1484 | -6 |
| 89 | | 12.0 | 97.0 | 1696 | 7 |
| 146 | | -11.0 | 95.0 | 1743 | 10 |

Average error 1.8% Standard Deviation 10% Median Miss Distance 7%

Table 1. Range estimates and percentage errors in calculated range for various transmitters, for propagation modes identified as normal E modes. Estimates were obtained using the one-dimensional quasi-parabolic model. Elevation angles flagged negative corresponded to secondary peaks in the elevation angle histograms. The mode is 1E, except for cases 45, 54 60, 87, 5, 89 and 146, for which it was 2E. The given apogee is then for the first hop.

| CASE | TX | APOGEE | [1]QP | [2]QP | DUD |
|---|---|---|---|---|---|
| 12 | NWCAPE | 191 | 4 | -6 | -3 |
| 42 | (2783km) | 197 | -11 | -11 | -6 |
| 88 | | 192 | -8 | -7 | -2 |
| 103 | | 189 | 1 | -1 | 2 |
| 147 | | 208 | -17 | -16 | -14 |
| 183 | | 197 | -10 | -9 | -3 |
| 38 | DARWIN | 193 | 2 | 7 | 11 |
| 100 | (2596km) | 193 | -8 | -7 | -6 |
| 120 | | 197 | -10 | -9 | -5 |
| 160 | | 212 | 1 | 4 | 6 |
| 186 | | 210 | -9 | -10 | -5 |
| 21 | BRISBANE | 207 | 2 | 1 | 11 |
| 44 | (1579km) | 197 | 13 | 10 | -14 |
| 52 | | 204 | 3 | 3 | 8 |
| 66 | | 199 | -9 | -7 | -2 |
| 137 | | 199 | -1 | -4 | 1 |
| 146 | | 212 | 0 | -1 | 3 |
| 165 | | 198 | -3 | -1 | 2 |
| 171 | | 229 | 7 | 9 | 16 |
| 204 | | 208 | -7 | -9 | -3 |
| 9 | TOWNSVILLE | 186 | -13 | -13 | -13 |
| 41 | (1893km) | 183 | -16 | -19 | -18 |
| 50 | | 194 | -13 | -16 | -18 |
| 92 | | 195 | -8 | -9 | -4 |
| 132 | | 168 | -8 | -11 | -11 |
| 144 | | 202 | -7 | -9 | -10 |
| 149 | | 200 | -6 | -6 | -7 |
| 162 | | 202 | -8 | -7 | -2 |
| 180 | | 175 | 5 | -7 | -5 |
| 157 | CANBERRA | 212 | 18 | 18 | 24 |
| 168 | (953km) | 199 | 0 | 7 | 1 |
| | AVERAGE % | | -3.7 | -4.4 | -2.1 |
| | SIGMA % | | 8.2 | 8.5 | 9.5 |
| | MEDIAN MISS DISTANCE % | | 7 | 5.5 | 7 |

Table 2. Percentage errors in range estimates for various transmitters, for propagation modes identified as F2 modes. Estimates are given for three models, with the apogee corresponding to the [1]QP model.

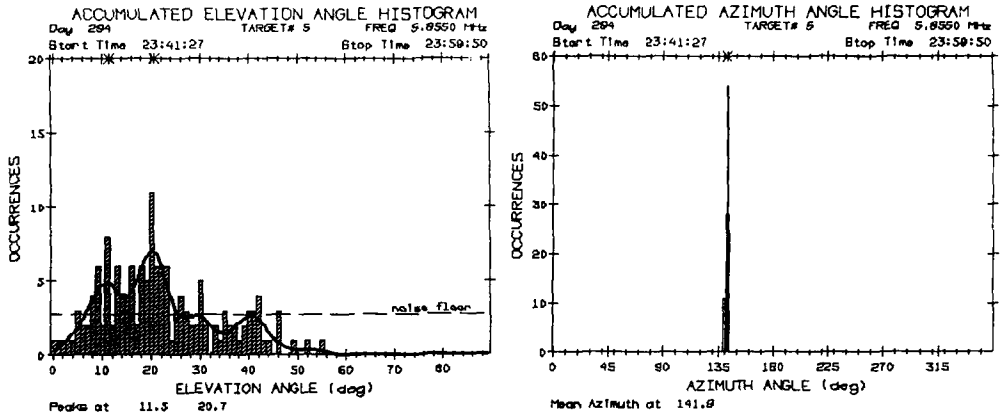



Fig 1. Accumulated histograms of elevation angle (left) and azimuth (right) for reception at Adelaide of a 5.855MHz signal from Hobart. The observations were accumulated for a nominal 20 minute period, from 23:41:27UT to 23:58:50UT. The curve through the elevation angle histogram was defined by an automatic peak-picking procedure.

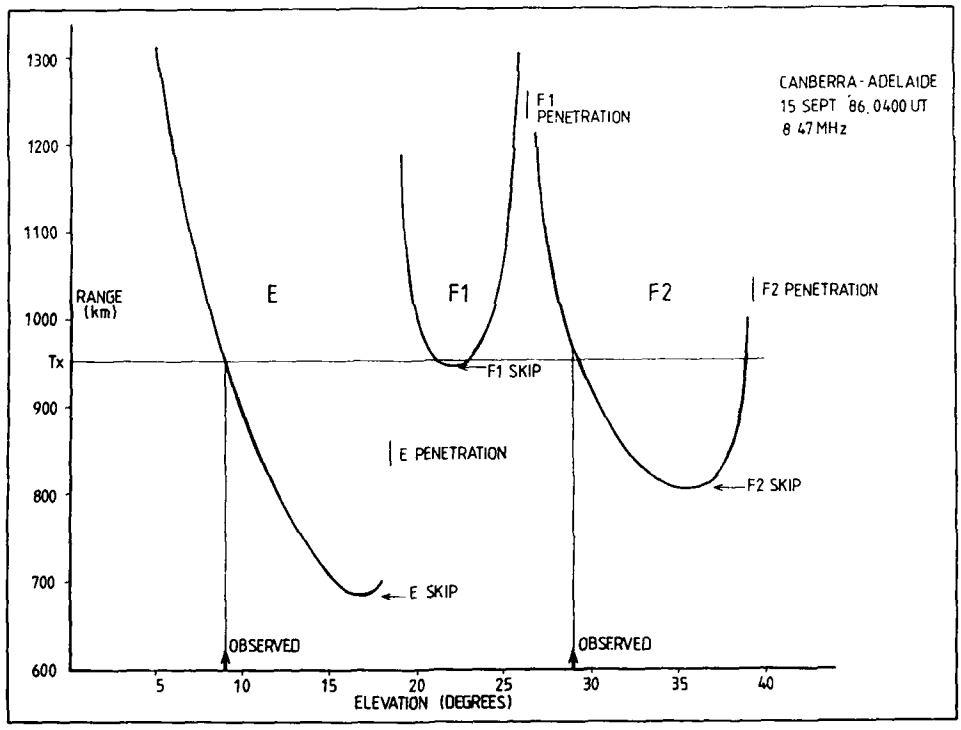



Fig 2. Calculated ground range as a function of elevation angle for an 8.47MHz signal propagating through a simple 3-layer model of the ionosphere, simulating the Canberra-Adelaide circuit at 0400UT on 15 September, 1986.

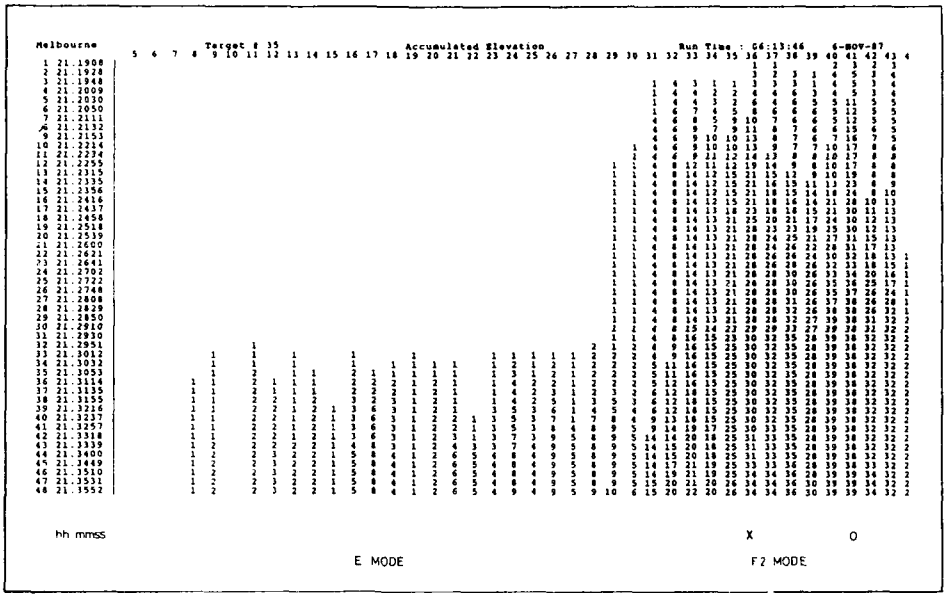



Fig 3. Accumulated numerical elevation angle histogram for a Melbourne-Adelaide circuit, showing a double-peaked F2 mode with early peaks at around 36 degrees and 41 degrees, apparently corresponding to X and O polarization modes. At the end of the accumulation period, there is only one unresolved peak, at 40.5 degrees. A mode transition occurred at around 21:30UT, the mode changing from 1F2 to 1E.

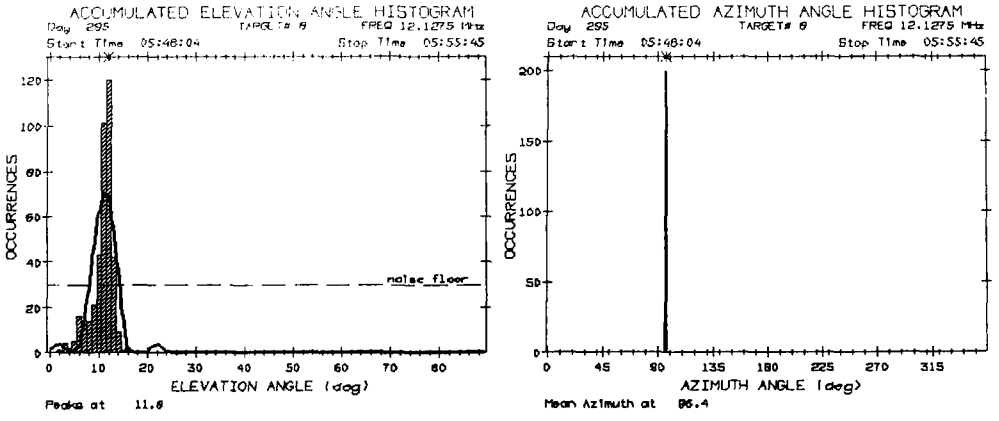



Fig 4. Accumulated histograms of elevation angle (left) and azimuth (right) for an Es propagation mode between Canberra and Adelaide. The calculated range is in excellent agreement with the correct range, and the mean bearing is only 0.1 degrees greater than the correct bearing of 96.3 degrees.

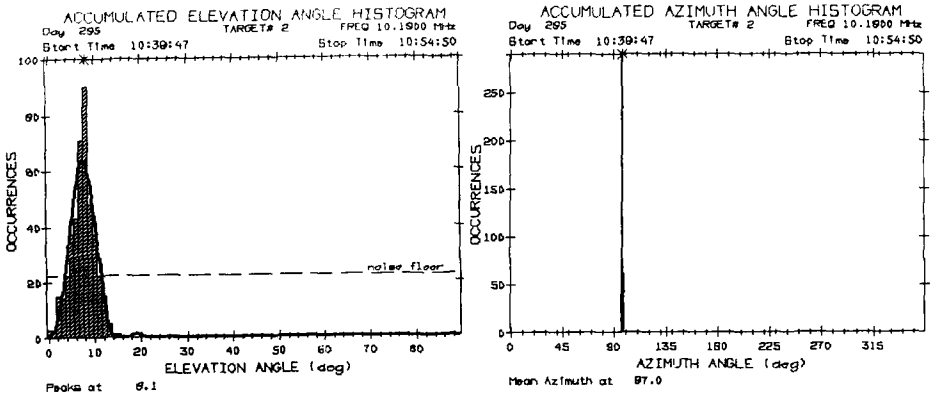



Fig 5. Accumulated histograms of elevation angle (left) and azimuth (right) for an Es propagation mode between Canberra and Adelaide. The calculated range is much greater than the correct range, and the mean bearing is 0.7 degrees greater than the correct bearing. The Es at the reflection point in this case seems to have been non-blanketing.

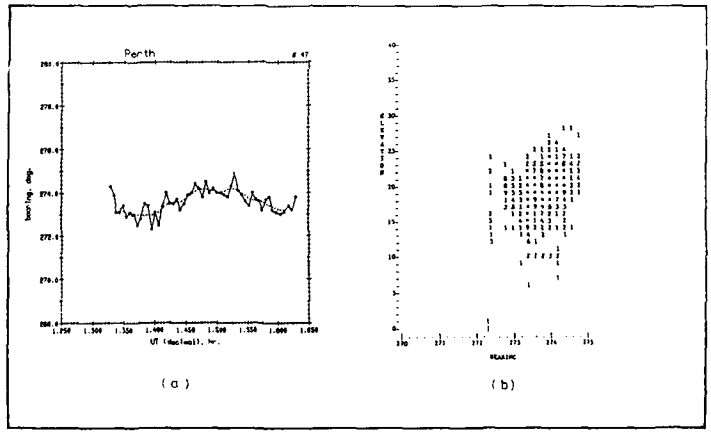



Fig 6. (a) The time series of the weighted mean bearing of a 10.540MHz signal from Perth being observed at Adelaide, showing the effects of a large TID.

(b) Numerical histograms of elevation versus bearing for the same data. Asterisks denote bin counts greater than 9, and the double vertical bar denotes the correct bearing.

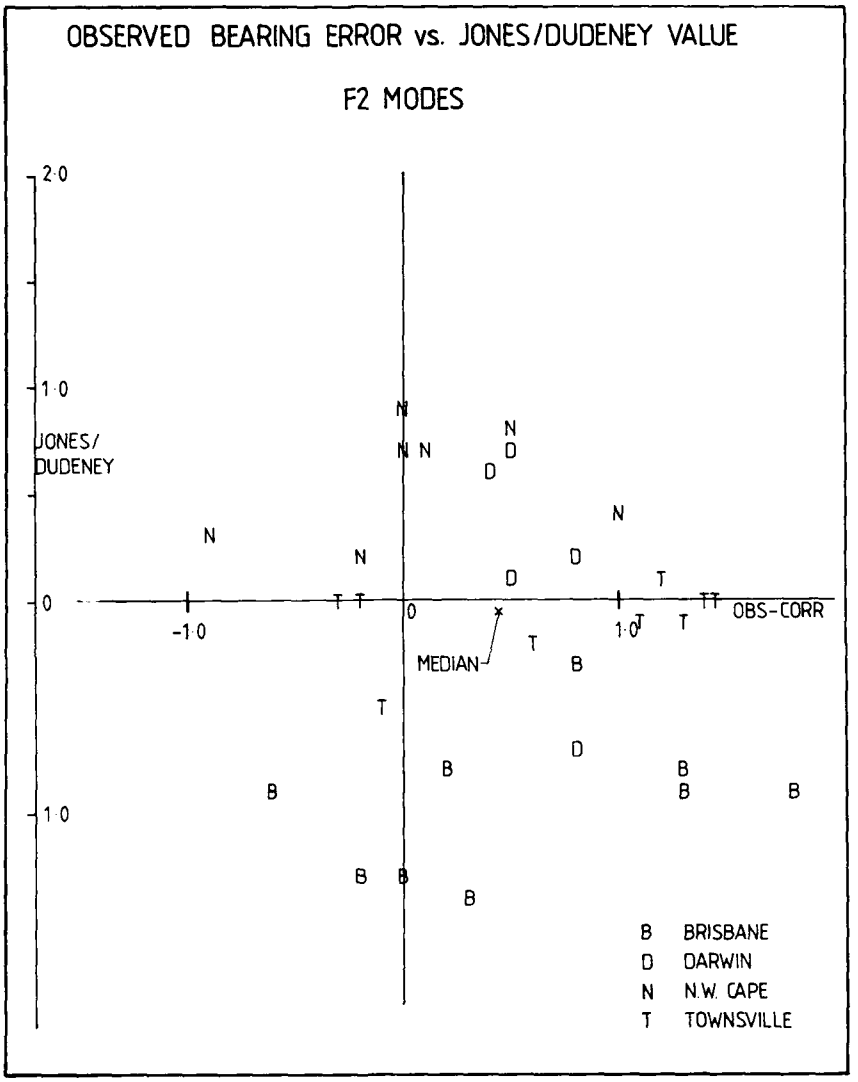



Fig 7. Bearing corrections obtained by 3-D raytracing, plotted against observed bearing errors, for nF2 modes from Brisbane, Darwin, Northwest Cape and Townsville to Adelaide. There appears to be no correlation between the two sets of results.

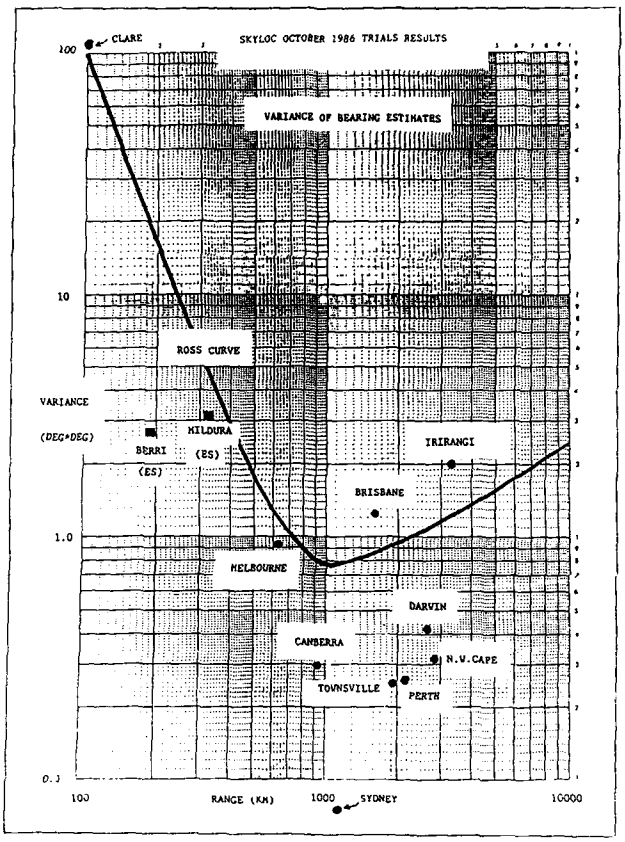



Fig 8. Bearing errors given by SKYLOC during the October 1986 trials, plotted in conjunction with the Ross curve. The Ross curve is an empirical curve of bearing variance versus target range, based on good quality DF data obtained over many circuit-years.

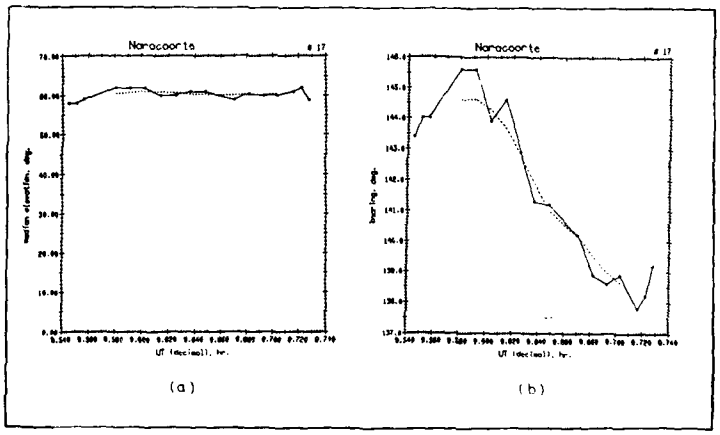



Fig 9. Time series of elevation angle and bearing for a 313km long circuit from Naracoorte to Adelaide. The series shows the effects of a large TID travelling at right angles to the circuit. The bearing swing is about 3 degrees either side of the correct bearing. The elevation angle stays remarkably constant and provides a range estimate with an error of around 10%.

## DISCUSSION

**T.Croft**

I compliment the author on an honest and candid report telling both the good and poor sides of present-day SSL. My own experience is in almost total agreement. I have two related comments:

(1) It is a convention in DF and SSL to quote median errors as the author does. I believe this practice has evolved because the resulting value is smaller than the mean would be. The long "tail" on the data distribution has little effect on the median, and so this parameter gives an optimistic final measure of quality. Yet, a text on statistics seldom mentions "median" because it has little relationship to the many other measures that typify the field of statistics.

(2) The author closed with an observation that hardware is a step ahead of ionospherics with regard to SSL. I agree, but with a single exception: SSL would derive considerable benefit if the receiving and processing system could, in the presence of multiple ionospheric modes, give up multiple correct arrival angles.

**Author's Reply**

I agree with your comments.

**K.Bibl**

Have you also made statistical studies of the range error for different transmitter locations, in addition to the statistic of the absolute amount of the range error?

**Author's Reply**

Yes. A paper describing the results of the October 1986 trials is to be published in JATP.

# INFLUENCE DES VARIATIONS IONOSPHERIQUES SUR LES SYSTEMES HF A HAUTE FIABILITE UTILISANT DE GRANDES BASES

C. GOUTELARD (1) et J.P. VAN UFFELEN (2)

(1) LETTI, 9 Avenue de la Division Leclerc, 94230 CACHAN, France

(2) TRT, 5 Avenue Réaumur, 92350 LE PLESSIS-ROBINSON, France

## 1 - INTRODUCTION

Les télécommunications par ondes décamétriques sont caractérisées par leur grande portée. L'encombrement spectral très important de cette gamme est un obstacle à une bonne fiabilité des liaisons qui, compte tenu de la variabilité du milieu, doivent être faites à fréquence variable. Il est souvent nécessaire d'obtenir des liaisons sûres avec des procédures ne requierant qu'un minimum de contrainte. La diversité d'espace est une technique bien connue. L'utilisation de trés grandes bases qui en est une extrapolation apporte des solutions aux formes spécifiques des problèmes liés à cette gamme : propagation et brouillage. Les télécommunications se font alors dans le concept de réseau de transmission dont la gestion optimale doit d'abord prendre en compte la modélisation à grande échelle de l'ionosphère. Les dimensions minimales de la base sont imposées par la décorrélation du brouillage observé aux stations de réception.

L'étude de ces deux problèmes a conduit à la définition de trois méthodes. Trois systèmes d'application hiérarchisés sont proposés.

## 2 - ANALYSE SPATIO-TEMPORELLE DU BRUIT

### 2.1. Principes des mesures

L'obtention d'une grande fiabilité de transmission nécessite le choix d'une fréquence permettant d'obtenir le rapport signal à bruit requis pour satisfaire l'utilisateur. Les premières analyses nous ont rapidement orientés vers l'utilisation de diversité de fréquence et d'espace. La détermination de la géométrie du réseau est basée en particulier sur les caractéristiques spatio-temporelle du bruit. Pour évaluer les propriétés statistiques du bruit, il a été procédé à une expérimentation ayant pour but de mesurer simultanément le niveau de bruit reçu dans diverses stations. Pour permettre l'analyse temporelle, ces enregistrements ont été synchronisés et datés avec une précision meilleure que la seconde. Les niveaux de bruit ont été mesurés dans 50 canaux de largeur de bande 5 KHz. Ces canaux répartis dans la gamme 3 à 30 MHz ont été choisis parmi ceux qui sont statistiquement calmes. Chaque canal est échantillonné simultanément sur l'ensemble des stations d'analyse, une fois par seconde. La répartition du niveau de bruit dépendant de la direction d'observation, nous avons utilisé des antennes directives.

Pour illustrer les résultats obtenus, nous divisons l'analyse d'une série de mesures effectuées dans trois stations et couvrant une période de 24 heures par tranche de deux heures, la direction d'observation étant le nord. Ces trois stations sont aux sommets d'un triangle dont les côtés ont pour longueur 1000, 700 et 500 Km.

### 2.2. Exploitation des mesures

Les mesures ont été exploitées tout d'abord indépendemment pour chaque station puis pour les trois couples et enfin sur l'ensemble des trois. Pour chaque station, à partir des mesures des niveaux de bruit S, on a déterminé pour un seuil donné Sc, choisi parmi 15 seuils distants de 3 dB (fig. 1) :

- Nc nombre de fois où $S < Sc$
- Dc la durée moyenne pendant laquelle $S < Sc$ $Dc = \frac{Tc}{Nc}$
- d écart type de la fonction de répartition des durées de clarté Tc .

L'exploitation simultanée des ensembles de 2 ou 3 stations a été faite selon les même principes. Les durées de clarté sont celles pour lesquelles les niveaux de bruit observés sur les 2 ou 3 stations considérées, sont simultanément inférieurs au seuil.

Ces exploitations faites par tranche de 2 heures permettent d'obtenir les durée moyennes de clarté et les écarts-types dont les notations sont indiquées dans le tableau 1.

## 2.3. Modèle statistique du bruit

### 2.3.1. Analyse par stations

A partir des valeurs Dc et d on a déterminé la densité de probabilité du niveau de bruit, c'est-à-dire la probabilité pour que le bruit soit compris entre deux valeurs de seuil séparés de 3 dB. Cette analyse a permis de constater la similitude de forme (loi log normale) de ces fonctions quelque soit la station, seules les valeurs moyennes du bruit observé sur une tranche de 2 heures diffèrent. Les figures 2.a,b,c donnent la fonction de répartition du bruit observé dans chaque station.

### 2.3.2. Analyse par couple de stations

Les performances obtenues par la diversité d'espace sont liées aux propriétés statistiques des bruits observés simultanément dans les diverses stations de réception. Ces stations ont été caractérisées en évaluant la corrélation du niveau de bruit pour les trois couples de stations.

Pour lesstations J et K, on a déterminé :

- les probabilités Pj (Sc) et Pk (Sc) d'apparition d'un niveau $S < Sc$ dans chaque station,
- la probabilité Pjk (Sc) d'apparition de niveaux simultanément inférieurs à Sc dans les deux stations.

On en a déduit, pour le seuil Sc, le coefficient d'indépendance Ijk(Sc) associé aux stations J et K.

Par définition, ce coefficient est égal à 1 lorque les variables aléatoires sont indépendantes. Les valeurs des coefficients d'indépendance en fonction du seuil sont données par la figure 3 qui montre que le coefficient d'indépendance est supérieur à 0,9 pour des niveaux de bruit supérieur à environ 100 dB.

## 2.4. Stratégie de choix de canal

Il est intéressant de déterminer une stratégie de choix de canal à partir des analyses faites sur le bruit. A cette fin, les performances obtenues ont été évaluées en sélectionnant les canaux pour lesquels la durée moyenne de clarté, sur une période de deux heures, est supérieure à 30 s pour un seuil de clarté Sc donné. Ces performances sont caractérisées par le pourcentage de canaux sélectionnés selon ce critère et dont la probabilité de clarté est supérieure à 99, 95, 90, 85, 80 ou 50 %. Ces résultats sont présentés par le figure 4.

A titre d'exemple, pour un seuil de - 102 db, la station 1 dispose de 27 % de canaux pour lequel le bruit est inférieur au seuil pendant 99 % du temps, 60 % de canaux clairs pendant 95 % du temps et 78 % de canaux clairs pendant 90 % du temps.

L'analyse des résultats sur l'ensemble des stations montre qu'en utilisant ce critère, 80 % des canaux sélectionnés sont clairs pendant 85 % du temps, ce qui met en évidence sa validité.

| Nombre de stations | Résultats de l'exploitation des mesures | | |
|---|---|---|---|
| 1 | Dc1<br>d1 | Dc2<br>d2 | Dc3<br>d3 |
| 2 | Dc1.2<br>d 1.2 | Dc1.3<br>d 1.3 | Dc2.3<br>d 2.3 |
| 3 | Dc1.2.3<br>d 1.2.3 | | |

TABLEAU 1

## 3 - MODELE IONOSPHERIQUE

### 3.1. Paramètres de l'études

L'étude ionosphérique a été faite à partir des hypothèses suivantes :

Les liaisons à établir sont localisées dans une zone de 7000 km de diamètre centrée sur le centre de gravité de la base.

L'étude a été faire en prenant en compte l'activité solaire, les effets journaliers et saisonniers et les variations spatiales des caractéristiques au milieu.

On a retenu pour cette étude, trois valeurs de l'indice d'activité solaire R = 50, 100 et 150.

Les périodes de test ont été limitées aux cas les plus significatifs en raison de la complexité des calculs. On s'est d'abord limité aux deux saisons extrêmes hiver et été. On a finalement retenu :

- juillet 12 H correspondant à l'angle zenithal le plus faible,
- décembre 8 H et 16 H correspondant aux gradients d'ionisation horizontaux maximum,
- juillet 20 H et décembre 20 H correspondant à des ionosphères nocturnes avec gradients d'ionisation horizontaux.

Le modèle ionosphérique de Bradley a été adopté pour étudier la propogation des ondes.

La zone étudiée est centrée sur l'Europe Occidentale.

### 3.2. Modes de propagation

Les modes de propagation à prendre en compte dépendent de la longueur de la liaison à assurer.

L'étude statistique étendue aux cas cités précédemment fait ressortir l'apparition des modes possibles pour les différentes longueurs de liaison. On a considéré qu'un mode était optimum s'il donnait le champ maximum Emax. On a alors défini une plage (appelée plage de réception optimale) dans laquelle le champs reçu E est tel que :

- E max - 10 dB < E < Emax

Lorsque plusieurs modes donnent des champs dont la valeur est comprise dans cette plage, on a considéré leur existence simultanée.

A titre d'exemple, pour une liaison de longueur D : 1500 km, sur 1586 cas traités, la statistique donne les résultats suivants :

- Mode 1 F optimal : 39 % des cas,
- Mode 2 F optimal : 9 % des cas,
- Mode 1E optimal : 5 % des cas,

Modes simultanés présents :

- Modes 1F et 2F.... : 26 % des cas,
- Modes 1F et 1 E... : 5 % des cas,
- Modes 1F, 2F et 1E : 14 % des cas,
- Autres modes..... : 2 % des cas.

Il est apparu à travers cette étude que les liaisons étaient essentiellement assurées par les modes 1E, 1F, 2F présents séparément ou simultanément.

Les modes mixtes 1E-1F ou multiples 2E, 3F apparaissent rarement, on a en conséquence, mené l'étude uniquement sur les trois modes principaux : 1F, 2F, 1E.

Les zones optimales, telles qu'elles ont été définies, sont de formes annulaires dont la dimension radiale est de l'ordre de 600 km. La détermination exacte ne peut être faire que par des calculs de champ précis fort longs.

On a évité cette difficulté en remarquant que pour la couche F, les distantes de bond variaient en moyenne de 300 km pour des incréments de fréquence d'émission égaux à $\frac{Fo.F2}{4}$. La zone où le champ est maximum à 10 dB près est donc constituée par 2 zones contigues.

On a donc déterminé les zones de réception optimales en calculant les distances de bond pour les fréquences :

$$Fn : Fo_{F2} \left(1 + \frac{n}{4}\right)$$

avec

n 1..... 10

On a conservé les mêmes fréquences pour les calculs des modes E pour les raisons suivantes :

- ce choix réduit l'étendue des zones optimales sensiblement dans le même rapport que celui des zones couvertes,

- les comparaisons des modes E et F sont plus faciles à faire, et notamment en ce qui concerne les occultations.

Les figures 5 .a et b montrent des exemples relatifs respectivement aux modes F et E.

Les mêmes fréquences utilisées pour le mode 2F ont donné des résultats dont la figure 5 .c est un exemple.

### 3.3. Détermination des zones de réception optimales

On a utilisé pour cette étude la terminologie suivante :

- Fv note la fréquence veillée par une station d'écoute,

- Ft note la fréquence de transmission,

- Fréquence optimale, celle assurant le meilleur bilan de transmission, défini précédemment pour une station d'écoute,

- Fréquence optimale simultanée celle qui est simultanément optimale pour deux, trois ou 4 stations d'écoute.

#### 3.3.1. Principe de détermination des zones de réception optimales par une station utilisant un sondeur à rétrodiffusion

On suppose que la station centrale est équipée d'un sondeur à rétrodiffusion permettant l'examen des conditions de propagation et d'une station de sondage oblique bistatique capable de transmettre avec les signaux de sondage des informations aux utilisateurs situés dans la zone couverte.

Pour une station d'écoute de la base, la zone optimale pour une fréquence Fn est délimitée par les deux courbes définies par (n-1) et (n+1).

Ainsi, si on considère les courbes (n-1) et (n+1) établies pour 2 stations (figure 6 ) on peut définir pour chaque la zone de fréquence optimale Fn qui permet de déterminer la zone de réception optimale simultanée pour les deux stations. Pour la fréquence Fn+1 on peut déterminer une zone identique et ainsi définir les bases d'une couverture. Pour des fréquences comprises entre Fn et Fn+1, on obtient des zones intermédiaires qui montrent que la zone totale de réception optimale pour deux stations est l'enveloppe des "pointes extérieures" des zones Fn, Fn+1 successives.

La figure 7 montre un exemple de détermination de zones de réception optimale pour deux stations distantes de 700 Km pour le mode 1F. Il apparait qu'une grande partie de la zone peut être couverte par une réception optimale simultanée pour les deux stations et deux zones ne peuvent être couvertes par une réception optimale que sur une seule station. Ces deux zones ont été appelées "zones de non réception optimale".

L'étude a été appliquée aux différentes configurations possibles en considérant les bases constituées par 2, 3 ou 4 stations d'écoute.

L'exemple présenté montre la détermination des zones de non réception optimale simultanées par deux stations.

On a considéré que la réception était optimale si au moins deux stations d'écoute de la base recevaient les signaux dans des conditions optimales, par l'un des modes de propagation 1F, 2F ou 1E.

#### 3.3.2. Principe de détermination des zones de réception optimales par une station utilisant un sondeur oblique bistatique

Dans cette hypothèse le choix de la fréquence de transmission est fait avec des signaux de sondages obliques bistatiques qui ne donnent de renseignements que sur une liaison. On considère que la station émettrice du sondeur oblique se trouve à proximité de la sation d'écoute 1.

Un utilisateur choisit donc sa fréquence de transmission en fonction de cette station. S'il se trouve sur la courbe n, il choisit la fréquence Fn optimale pour la station 1. La réception est donc également optimale pour la station 2 entre les courbes n-1 et n+1 (figure 8 ), donc sur le segment AB. Pour les fréquences Fn+1, Fn+2 on détermine de même les segments A, B1, A2, B2.

La zone optimale simultanée est donc limitée par les lieux des points A et B.

On peut remarquer alors que les zones optimales de réception sont moins étendues dans ce cas que dans celui où l'on utilise un sondeur à rétrodiffusion. En effet, il apparait sur la figure 9 que :

- la zone optimale dans le cas d'une station équipée d'un sondeur à rétrodiffusion est obtenue par l'enveloppe des "pointes extérieures" des courbes (n-1), (n+1) .

- la zone optimale dans le cas d'une station équipée uniquement d'un sondeur oblique bistatique est obtenue pour l'enveloppe des "pointes internes" des courbes (n-1), (n+1) .

La figure 7 fait apparaître la différence des zones de réception non optimales pour les deux cas. L'étude a été appliquée aux différentes configurations de bases équipées de 2, 3 ou 4 stations d'écoute.

La procédure devient identique à celle décrite précédemment en utilisant les zones de non réception optimales simultanées.

## 3.4. Résultats

### 3.4.1. Utilisation du sondage oblique bistatique

Dans le cas d'une base constituée de deux stations d'écoute, la couverture est médiocre.

Dans le cas d'une base définie au paragraphe 2.1. et constituée de 3 stations d'écoute, les résul-sultats sont donnés dans le tableau 3.3. pour tous les indices solaires confondus.

| | JUILLET 12 H | | JUILLET 20 H | | DECEMBRE 8 H | | DECEMBRE 16 H | | DECEMBRE 20 H | | MOYENNE SAISONS | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | S | % | S | % | S | % | S | % | S | % | S | % |
| 1E | 1,29 $10^6$ | 8,90 | - | - | 1,59 $10^6$ | 4,1 | 2,28 $10^6$ | 5,7 | - | - | 1,68 $10^6$ | 9,38 |
| 1F | 7,15 $10^6$ | 18,23 | 5,60 $10^6$ | 14,28 | 5,71 $10^6$ | 14,56 | 7,28 $10^6$ | 18,56 | 4,54 $10^6$ | 11,58 | 6,06 $10^6$ | 15,45 |
| 2F | 0 | 0 | 0 | 0 | 4,14 $10^6$ | 10,55 | 2,11 $10^6$ | 5,38 | 0 | 0 | 1,12 $10^6$ | 7,97 |
| T* | 2,47 $10^6$ | 6,30 | 5,60 $10^6$ | 14,28 | 4,03 $10^6$ | 10,28 | 3,43 $10^6$ | 8,75 | 4,31 $10^6$ | 11,14 | 3,38 $10^6$ | 10,15 |

* T : tous modes confondus sondage oblique bistatique

TABLEAU 2

ZONES DE NON RECEPTION OPTIMALE SIMULTANEE

Une représentation géographique des résultats est faite sur la figure 10 qui montre pour tous les cas étudiés et par l'utilisation des 3 modes possibles, la probabilité d'avoir une réception optimale.

Dans le cas d'une base constituée pour 4 stations d'écoute, les résultats sont résumés dans le tableau 3 pour tous les cas et tous modes confondus.

| JUILLET 12 H | | JUILLET 20 H | | DECEMBRE 8 H | | DECEMBRE 16 H | | DECEMBRE 20 H | | MOYENNES SAISONS | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| S | % | S | % | S | % | S | % | S | % | S | % |
| 1,29 $10^6$ | 3,29 | 3,06 $10^6$ | 7,80 | 1,90 $10^6$ | 4,82 | 7,22 $10^5$ | 84 | 2,58 $10^6$ | 6,59 | 1,91 $10^6$ | 4,87 |

TABLEAU 3

La représentation géographique des résultats est donnée sur la figure 11.

3.4.2. Utilisation du sondage par rétrodiffusion

Pour une base constituée par deux stations d'écoute la couverture est médiocre. Dans le cas de la base constituée de 3 stations d'écoute, les résultats sont donnés dans le tableau 4, pour tous indices solaires confondus.

| | JUILLET 12 H | | JUILLET 20 H | | DECEMBRE 8 H | | DECEMBRE 16 H | | DECEMBRE 20 H | | MOYENNE SAISONS | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | S | % | S | % | S | % | S | % | S | % | S | % |
| 1E | $9,64\ 10^5$ | 2,46 | - | - | $1,49\ 10^6$ | 3,80 | $8,89\ 10^5$ | 2,27 | - | - | $1,11\ 10^6$ | 2,83 |
| 1F | $4,03\ 10^6$ | 10,28 | $7,22\ 10^5$ | 1,84 | $1\ 10^6$ | 2,55 | $1,95\ 10^6$ | 4,97 | $9,92\ 10^5$ | 2,53 | $1,74\ 10^6$ | 4,44 |
| 2F | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| T* | $6,94\ 10^4$ | 0,18 | $7,22\ 10^5$ | 1,84 | $3,72\ 10^5$ | 0,95 | $2,19\ 10^5$ | 0,56 | $9,92\ 10^5$ | 2,53 | $5,15\ 10^5$ | 1,31 |

* T : tous modes confondus sondage par rétrodiffusion.

TABLEAU 4

ZONES DE NON RECEPTION OPTIMALE SIMULTANEE

Une représentation géographique des résultats est faite sur la figure 12 qui représente pour tous les cas étudiés et par l'utilisation des 3 modes possibles, la probabilité d'avoir une réception optimale simultanée.

Dans le cas d'une base constituée par 4 stations d'écoute, la réception devient optimale sur toute la zone.

3.5. Conclusion sur les conditions de couverture de la zone

Les résultats obtenus font apparaître que :

- les performances de bonne couverture nécessitent des bases de 3 ou 4 stations d'écoute,
- l'utilisation d'un sondeur à rétrodiffusion améliore d'un facteur l'ordre de 7 la probabilité de bonne couverture,
- l'influence de l'activité solaire ne paraît pas importante,
- les couvertures les meilleures sont obtenues l'été à 12 heures, au moment où les gradients horizontaux d'ionisation sont minimaux,
- les couvertures les plus mauvaises sont obtenues la nuit à cause de l'absence de la couche E,
- la position des stations d'écoute est importante dans le sens où elle localise les zones de mauvaise réception.

4 - DESCRIPTION DES SYSTEMES PROPOSES

L'étude présentée visait la conception de systèmes à très haute fiabilité garantissant cependant aux utilisateurs une souplesse maximale, par l'utilisation optimale des caractéristiques du canal ionosphériques évaluées sur une grande échelle.

4.1. Hypothèses

Les systèmes proposés sont conçus avec comme objectif l'obtention d'une transmission à haute fiabilité d'un ou plusieurs messages émis sans contrainte de préavis ou de rendez-vous entre les correspondants. Les destinataires ne connaissent pas la localisation de l'émetteur.

Pour être efficace, le processus de choix d'une fréquence de transmission doit tenir compte des conditions de propagation et du bruit radioélectrique au lieu de réception. Trois systèmes ont été proposés pour répondre à l'objectif ; ils ont tous en commun l'utilisation de la diversité d'espace et de fréquence. Un ensemble de N fréquences est veillé en permanence par M récepteurs (M > N) répartis en 2, 3 ou 4 stations.

Par contre, ils diffèrent par le méthodes d'évaluation des conditions de propagation en temps réel (directes ou indirectes) d'une part et du niveau de bruit local (en temps réel ou par prévisions) d'autre part.

4.2. Système I

4.2.1. Principe

Ce système est basé sur :

- l'utilisation d'un sondage oblique bistatique pour évaluer en temps réel les conditions de propagation,
- la connaissance par l'émetteur de la liste des fréquences de veille établie au préalable,
- le choix par l'émetteur de la fréquence de transmission dans cette liste par analyse de l'ionogramme reçu en temps réel.

4.2.2. Sélection des fréquences veillées

Les listes des fréquences veillées sont déterminées à l'aide des prévisions ionosphériques à long terme et en tenant compte du niveau de bruit sur les fréquences utilisables évalué par une analyse statistique du bruit radioélectrique aux lieux de réception.

Les fréquences veillées sont donc déterminées de façon doublement statistique.

4.3. Système II

4.3.1. Principe

Ce système diffère du précédent par le fait que le choix des fréquences veillées est effectué en temps réel par l'analyse permanente, dans les stations d'écoute, du bruit local pour déterminer parmi les fréquences choisies au vu des prévisions ionosphériques à long terme, celles qui sont les moins brouillées.

Les fréquences veillées sont communiquées à l'utilisateur grâce à l'adjonction d'une capacité de transmission de données au sondeur oblique bistatique.

4.3.2. Méthode du choix des fréquences veillées

La détermination de la plage de fréquence assurant la couverture se fait comme avec le système I, par l'utilisation des prévisions ionosphériques à long terme.

Le choix des fréquences veillées est effectué en temps réel à l'intérieur de cette plage. Chaque station mesure le niveau de bruit affectant tous les canaux de cette plage et les transmet à une station dite centrale qui effectue le choix à partir des critères suivants :

- niveau instantané du bruit, dans les stations d'écoute,
- historique de l'occupation du canal considéré.

La liste des fréquences veillées est communiquée aux utilisateurs à l'aide des signaux d'information émis simultanément avec les signaux de sondage, et directement aux stations d'écoute.

4.3.3. Choix de la fréquence de transmission

Le choix de la fréquence de transmission est effectué, par l'utilisateur, comme avec le système I par l'analyse de l'ionogramme oblique.

Le choix de la fréquence est donnée automatiquement par les informations accompagnant les signaux de sondage.

4.3.4. Conclusion

La méthode utilisant une liaison rest avec prise en compte en temps réel du bruit aux lieux de réception est une méthode :

- adaptative, fonctionnant en temps réel en ce qui concerne le brouillage,
- adaptive dans une gamme définie de façon prévisionnelle en ce qui concerne le choix des fréquences veillées,
- en temps réel en ce qui concerne la fréquence de transmission.

## 4.4. Système III

### 4.4.1. Principe

Cette méthode consiste à adjoindre à la précédente un dispositif rendant adaptatif le choix de la fréquence de transmission.

Le principe consiste à utiliser une station de sondage par rétrodiffusion pour dresser une "cartographie de l'ionosphère" permettant de choisir au mieux les fréquences optimales en minimisant les zones de non réception simultanée par 2 (3 ou 4) stations.

Le choix des fréquences veillées peut alors être effectué comme dans la seconde méthode avec transmission des informations aux stations d'écoute et aux utilisateurs.

La station centrale ayant la connaissance des conditions de propagation et des niveaux de bruit, gère les fréquences de transmission de façon totalement adaptative en fixant ces fréquences aux stations d'écoute et aux utilisateurs selon le procédé décrit ci-dessous.

### 4.4.2. Choix des fréquences veillées

La station de sondage par rétrodiffusion établit la "cartographie de l'ionosphère" et détermine les zones de réception optimale simultanée pour étudier la couverture de la zone.

Cette détermination est faite de la façon suivante :

On définit des zones annulaires à partir des conditions de propagation de façon qu'à l'intérieur d'une de ces zones il existe une fréquence de réception optimale simultanée pour au moins deux stations d'écoute. Un calcul statistique a montré que cette détermination se fait pour 97,85 % de la surface de la zone. Pour les 2,15 % restant la réception ne peut être optimale que sur l'une seule station d'écoute. La station centrale détermine pour ces zones la fréquence donnant le champ maximum dans l'une ou l'autre des deux stations ou la réception n'est pas optimale.

La détermination des fréquences veillées est faite par la station centrale où se trouvent rassemblées les données relatives à la propagation (zones annulaires) et les résultats de l'analyse du bruit transmis comme dans le cas du système II. La détermination se fait séparément pour chaque zone annulaire.

Pour une zone annulaire, la station centrale connait les deux stations d'écoute pour lesquelles les réceptions sont optimales et pour chacune de ces stations elle connait :

- le niveau du champ reçu pour la fréquence optimale,
- les fréquences dites claires immédiatement voisines et leur niveau de clarté.

Ces deux données permettent de calculer pour chaque fréquence claire la probabilité d'erreur pour chaque station d'écoute et par conséquent la probabilité de bonne réception par les deux stations d'écoute.

La fréquence de transmission est celle qui parmi les fréquences claires immédiatement voisines de la fréquence optimale, donne la meilleur probabilité de bonne réception.

### 4.4.3. Transmission de la fréquence veillée à l'utilisateur

La fréquence de transmission est assignée à l'utilisateur par l'intermédiaire du sondeur qui fait une bistatique où les signaux d'information sont ajoutés aux signaux de sondage comme dans le cas du système II.

### 4.4.4. Conclusion

La méthode utilisant un sondeur à rétrodiffusion est une méthode :

- adaptative, fonctionnant en temps réel en ce qui concerne le brouillage,
- adaptative, fonctionnant en temps réel en ce qui concerne le choix des fréquences veillées,
- adaptative, fonctionnant en temps réel en ce qui concerne le choix de la fréquence de trafic.

## 5 - CONCLUSION

La fiabilité des transmissions dépend des paramètres de la propagation et du brouillage. L'utilisation de grandes bases apporte des solutions à ce problème. Leur dimension minimale est fixée par les brouillages. La gestion du réseau formé nécessite la modélisation sur une grande échelle du milieu de propagation et l'évaluation du brouillage.

On a montré dans quelles conditions une réception optimale pouvait être obtenue dans le cas de bases comportant de 2 à 4 stations d'écoute. Les couvertures optimales limites on été déterminées.

Trois systèmes hiérarchisés ont été définis par une gestion en temps réel plus ou moins importante. Dans le cas d'une gestion totalement en temps réel, la couverture optimale peut toujours être obtenue en dehors d'absorption exceptionnelle type PIDB. De plus, la hiérarchie des systèmes permet dans un mode opérationnel d'envisager des fonctionnements en mode dégradé.

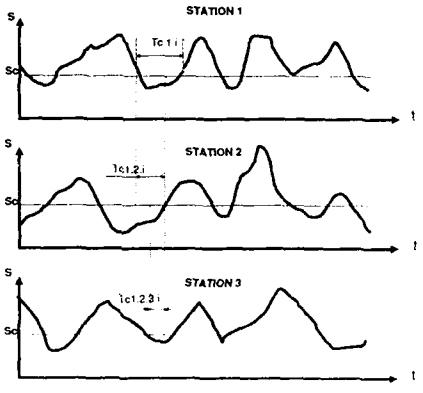
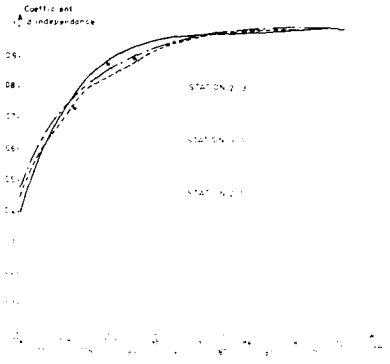



FIGURE 1

ANALYSE DU BRUIT



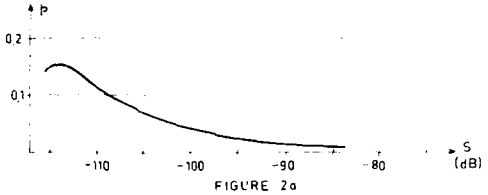



FIGURE 2a

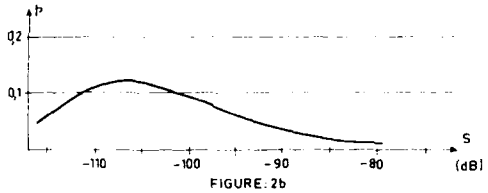



FIGURE 2b

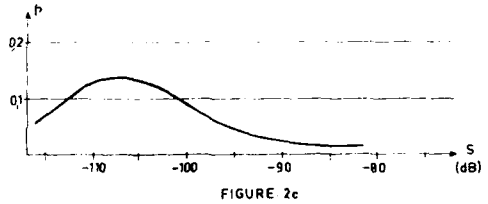
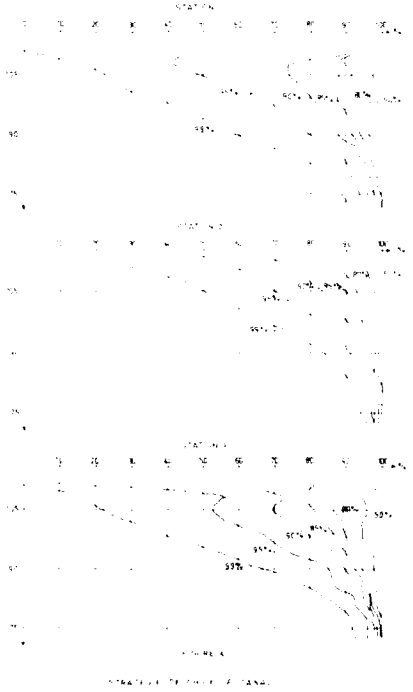



FIGURE 2c

DENSITE DE PROBABILITE DU BRUIT



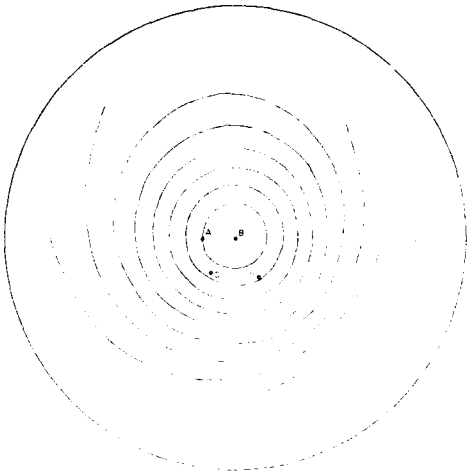


FIGURE 5a
Zones de réception optimale
Juillet_12h_IR=50_Mode 1F

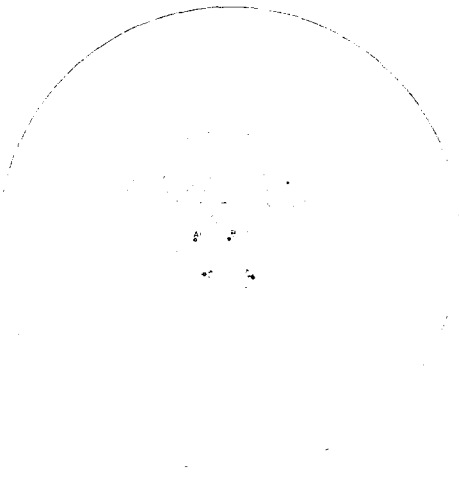


FIGURE 5b
Zones de réception optimale
Juillet_12h_IR=50_Mode 1 E



FIGURE 5c
Zones de réception optimale
Juillet_12h_IR=50_Mode 2 F

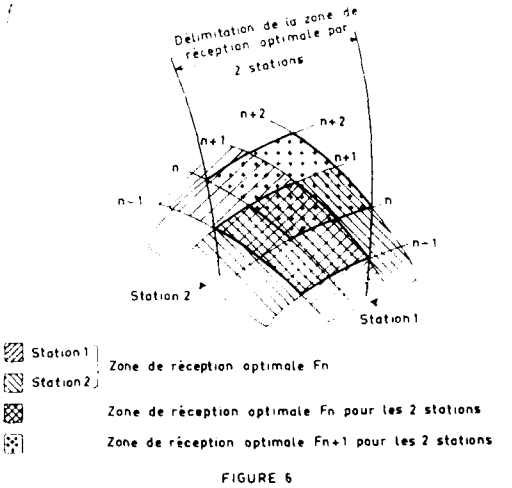



FIGURE 6
METHODE DE DETERMINATION DES ZONES DE RECEPTION OPTIMALES
SONDAGE PAR RETRODIFFUSION

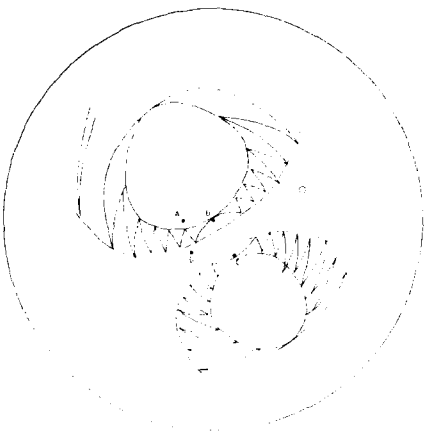


FIGURE 7
ZONES ELEMENTAIRES DE RECEPTION SIMULTANEE
JUILLET 12h IR=50 MODE 1F

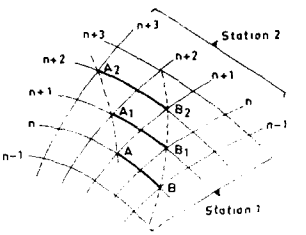
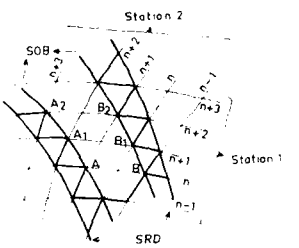



FIGURE 8
METHODE DE DETERMINATION
DES ZONES DE RECEPTION OPTIMALES
SONDAGE OBLIQUE BISTATIQUE




SOB Limites de la zone avec une seule station de sondage de ligne bistatique

SRD Limites de la zone avec une station centralisee

FIGURE 9
COMPARAISON DES METHODES DE DETERMINATION
DE RECEPTION OPTIMALE

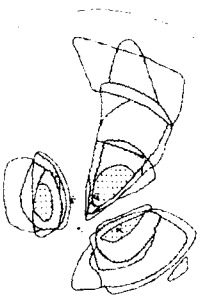


FIGURE 10
RESULTATS SUMMUM
SONDAGE OBLIQUE BISTATIQUE 3 STATIONS

INDICATION DES LEGENDES
NOMBRE DE FOIS OU LA RECEPTION
N'EST PAS OPTIMALE

0 FOIS
1 FOIS
2 FOIS
3 FOIS
4 FOIS
5 FOIS

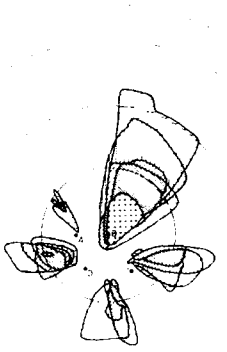


FIGURE 11
RESULTATS CUMULES
SONDAGE OBLIQUE BISTATIQUE 4 STATIONS

INDICATION DES LEGENDES
NOMBRE DE FOIS OU LA RECEPTION
N'EST PAS OPTIMALE

0 FOIS
1 FOIS
2 FOIS
3 FOIS
4 FOIS
5 FOIS



FIGURE 12
RESULTATS CUMULES
SONDAGE PAR RETRODIFFUSION 3 STATIONS

INDICATION DES LEGENDES
NOMBRE DE FOIS OU LA RECEPTION
N'EST PAS OPTIMALE

0 FOIS
1 FOIS
2 FOIS
3 FOIS
4 FOIS
5 FOIS

## REPORT DOCUMENTATION PAGE

| 1. Recipient's Reference | 2. Originator's Reference | 3. Further Reference | 4. Security Classification of Document |
|---|---|---|---|
| | AGARD-CP-441 | ISBN 92-835-0500-X | UNCLASSIFIED |

| | |
|---|---|
| **5. Originator** | Advisory Group for Aerospace Research and Development<br>North Atlantic Treaty Organization<br>7 rue Ancelle, 92200 Neuilly sur Seine, France |
| **6. Title** | IONOSPHERIC STRUCTURE AND VARIABILITY ON A GLOBAL SCALE AND INTERACTIONS WITH ATMOSPHERE AND MAGNETOSPHERE |
| **7. Presented at** | the Electromagnetic Wave Propagation Panel Symposium held in Munich, Germany, 16—20 May 1988 |

| 8. Author(s)/Editor(s) | 9. Date |
|---|---|
| Various | April 1989 |
| **10. Author's/Editor's Address** | **11. Pages** |
| Various | 498 |

**12. Distribution Statement** This document is distributed in accordance with AGARD policies and regulations, which are outlined on the Outside Back Covers of all AGARD publications.

**13. Keywords/Descriptors**

| | |
|---|---|
| *Ionospheric structure behaviour* | Communication |
| Solar wind | Global interaction |
| Propagation media | |

**14. Abstract**

Communications, Navigation and Surveillance Systems operating in, through the aerospace EM propagation environment are affected by the state variability of the propagation media. The range of phenomena need for their elucidation, observations and analysis on a global scale since only an understanding of the complex global interaction can improve the means of predictability and assessment of localised phenomena suggesting methods for mitigating adverse propagation conditions.

With this goal, ionospheric dynamics, ionosphere magnetosphere and ionosphere atmosphere interactions were analysed and discussed at this symposium.

| AGARD Conference Proceedings No.441<br>Advisory Group for Aerospace Research and Development, NATO<br>IONOSPHERIC STRUCTURE AND VARIABILITY ON A GLOBAL SCALE AND INTERACTIONS WITH ATMOSPHERE AND MAGNETOSPHERE<br>Published April 1989<br>498 pages<br><br>Communications, Navigation and Surveillance Systems operating in/through the aerospace EM propagation environment are affected by the state/variability of the propagation media. The range of phenomena need for their elucidation, observations and analysis on a global scale since only an understanding of the complex global interaction can improve the means of predictability and<br><br>P.T.O. | AGARD-CP-441<br><br>Ionospheric structure/ behaviour<br>Solar wind<br>Propagation media<br>Communication<br>Global interaction |
|---|---|
| AGARD Conference Proceedings No.441<br>Advisory Group for Aerospace Research and Development, NATO<br>IONOSPHERIC STRUCTURE AND VARIABILITY ON A GLOBAL SCALE AND INTERACTIONS WITH ATMOSPHERE AND MAGNETOSPHERE<br>Published April 1989<br>498 pages<br><br>Communications, Navigation and Surveillance Systems operating in/through the aerospace EM propagation environment are affected by the state/variability of the propagation media. The range of phenomena need for their elucidation, observations and analysis on a global scale since only an understanding of the complex global interaction can improve the means of predictability and<br><br>P.T.O. | AGARD-CP-441<br><br>Ionospheric structure/ behaviour<br>Solar wind<br>Propagation media<br>Communication<br>Global interaction |



| AGARD Conference Proceedings No.441<br>Advisory Group for Aerospace Research and Development, NATO<br>IONOSPHERIC STRUCTURE AND VARIABILITY ON A GLOBAL SCALE AND INTERACTIONS WITH ATMOSPHERE AND MAGNETOSPHERE<br>Published April 1989<br>498 pages<br><br>Communications, Navigation and Surveillance Systems operating in/through the aerospace EM propagation environment are affected by the state/variability of the propagation media. The range of phenomena need for their elucidation, observations and analysis on a global scale since only an understanding of the complex global interaction can improve the means of predictability and<br><br>P.T.O. | AGARD-CP-441<br><br>Ionospheric structure/ behaviour<br>Solar wind<br>Propagation media<br>Communication<br>Global interaction |
|---|---|
| AGARD Conference Proceedings No.441<br>Advisory Group for Aerospace Research and Development, NATO<br>IONOSPHERIC STRUCTURE AND VARIABILITY ON A GLOBAL SCALE AND INTERACTIONS WITH ATMOSPHERE AND MAGNETOSPHERE<br>Published April 1989<br>498 pages<br><br>Communications, Navigation and Surveillance Systems operating in/through the aerospace EM propagation environment are affected by the state/variability of the propagation media. The range of phenomena need for their elucidation, observations and analysis on a global scale since only an understanding of the complex global interaction can improve the means of predictability and<br><br>P.T.O. | AGARD-CP-441<br><br>Ionospheric structure/ behaviour<br>Solar wind<br>Propagation media<br>Communication<br>Global interaction |

assessment of localised phenomena suggesting methods for mitigating adverse propagation conditions.

With this goal, ionospheric dynamics, ionosphere/magnetosphere and ionosphere/atmosphere interactions were analysed and discussed at this symposium.

Papers presented at the Electromagnetic Wave Propagation Panel Symposium held in Munich, Germany, 16–20 May 1988.

ISBN 92-835-0500-X

assessment of localised phenomena suggesting methods for mitigating adverse propagation conditions.

With this goal, ionospheric dynamics, ionosphere/magnetosphere and ionosphere/atmosphere interactions were analysed and discussed at this symposium.

Papers presented at the Electromagnetic Wave Propagation Panel Symposium held in Munich, Germany, 16–20 May 1988.

ISBN 92-835-0500-X

assessment of localised phenomena suggesting methods for mitigating adverse propagation conditions.

With this goal, ionospheric dynamics, ionosphere/magnetosphere and ionosphere/atmosphere interactions were analysed and discussed at this symposium.

Papers presented at the Electromagnetic Wave Propagation Panel Symposium held in Munich, Germany, 16–20 May 1988.

ISBN 92-835-0500-X

assessment of localised phenomena suggesting methods for mitigating adverse propagation conditions.

With this goal, ionospheric dynamics, ionosphere/magnetosphere and ionosphere/atmosphere interactions were analysed and discussed at this symposium.

Papers presented at the Electromagnetic Wave Propagation Panel Symposium held in Munich, Germany, 16–20 May 1988.

ISBN 92-835-0500-X

AGARD
NATO OTAN
7 rue Ancelle · 92200 NEUILLY-SUR-SEINE
FRANCE
Telephone (1)47.38.57.00 · Telex 610 176

**DISTRIBUTION OF UNCLASSIFIED AGARD PUBLICATIONS**

AGARD does NOT hold stocks of AGARD publications at the above address for general distribution. Initial distribution of AGARD publications is made to AGARD Member Nations through the following National Distribution Centres. Further copies are sometimes available from these Centres, but if not may be purchased in Microfiche or Photocopy form from the Purchase Agencies listed below.

NATIONAL DISTRIBUTION CENTRES

BELGIUM
Coordonnateur AGARD — VSL
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evere, 1140 Bruxelles

CANADA
Director Scientific Ir
Dept of National De
Ottawa, Ontario K1

DENMARK
Danish Defence Res
Ved Idraetsparken
2100 Copenhagen

FRANCE
O.N.E.R.A. (Direct
29 Avenue de la Di
92320 Châtillon

GERMANY
Fachinformationsz
Physik, Mathemati
Karlsruhe
D-7514 Eggenstein

GREECE
Hellenic Air Force
*Aircraft Support E*
Department of Res
Holargos, Athens, T

ICELAND
Director of Aviation
c/o Flugrad
Reykjavik

ITALY
Aeronautica Militare
Ufficio del Delegato Nazionale all'AGARD
3 Piazzale Adenauer
00144 Roma EUR

LUXEMBOURG
*See* Belgium

NETHERLANDS
Netherlands Delegation to AGARD
National Aerospace Laboratory, NLR
P.O. Box 126
2600 AC Delft

UNITED KINGDOM
Defence Research Information Centre
Kentigern House
65 Brown Street
Glasgow G2 8EX

UNITED STATES
National Aeronautics and Space Administration (NASA)
Langley Research Center
M/S 180
Hampton, Virginia 23665

THE UNITED STATES NATIONAL DISTRIBUTION CENTRE (NASA) DOES NOT HOLD STOCKS OF AGARD PUBLICATIONS, AND APPLICATIONS FOR COPIES SHOULD BE MADE DIRECT TO THE NATIONAL TECHNICAL INFORMATION SERVICE (NTIS) AT THE ADDRESS BELOW.

PURCHASE AGENCIES

National Technical
Information Service (NTIS)
5285 Port Royal Road
Springfield
Virginia 22161, USA

ESA/Information Retrieval Service
European Space Agency
10, rue Mario Nikis
75015 Paris, France

The British Library
Document Supply Division
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
England

Requests for microfiche or photocopies of AGARD documents should include the AGARD serial number, title, author or editor, and publication date. Requests to NTIS should include the NASA accession report number. Full bibliographical references and abstracts of AGARD publications are given in the following journals:

Scientific and Technical Aerospace Reports (STAR)
published by NASA Scientific and Technical
Information Branch
NASA Headquarters (NIT-40)
Washington D.C. 20546, USA

Government Reports Announcements (GRA)
published by the National Technical
Information Services, Springfield
Virginia 22161, USA

*Printed by Specialised Printing Services Limited*
*40 Chigwell Lane, Loughton, Essex IG10 3TZ*

ISBN 92-835-0500-X

NASA
National Aeronautics and Space Administration
Washington, D.C. 20546
SPECIAL FOURTH CLASS MAIL BOOK
Postage and Fees Paid
National Aeronautics and Space Administration
NASA-451
Official Business
Penalty for Private Use $300

L2 001 AGARDCP450N9006143002/7EA
DEPT OF DEFENSE
DEFENSE TECHNICAL INFORMATION CENTER
DTIC-FDAC
CAMERON STATION BLDG 5
ALEXANDRIA VA 22304-6145

END

DATE
FILMED

9-89