ANARE REPORTS
146

Australian solar terrestrial and space physics research

Edited by R.J. Morris and P.J. Wilkinson

Australian
National
Antarctic
Research
Expeditions

ANARE Reports and *ANARE Research Notes*

The Australian Antarctic Division produces the two serials *ANARE Reports* and *ANARE Research Notes*. Anyone who has participated in research that was supported by the AAD is invited to publish through these series. Access to electronic publications is free through the Australian Antarctic Division website at www.aad.gov.au

The *ANARE Reports* series is peer reviewed, which is appropriate for manuscripts that are considered too lengthy or detailed for journal publication.

The *ANARE Research Notes* series is not peer reviewed, allowing for more rapid publication. This is appropriate for manuscripts that have a practical field application such as manuals, identification keys datasets and atlases.

Published May 2001 by
Australian Antarctic Division
Department of the Environment and Heritage
Channel Highway
Kingston
Tasmania 7050
Australia
email: publications@aad.gov.au

ISBN: 1 876934 00 10

AUSTRALIAN SOLAR TERRESTRIAL AND SPACE PHYSICS RESEARCH

Edited by

R.J. Morris[1] and P.J. Wilkinson[2]

[1]Australian Antarctic Division
Department of the Environment and Heritage
Kingston Tasmania 7050 Australia

[2]IPS Radio and Space Services
Department of Industry Science and Resources
Haymarket NSW 1240 Australia

PREFACE

The papers in this collection of middle and upper atmospheric physics research were presented at two recent Australian S-RAMP meetings:

(i) 13th Australian Institute of Physics Congress
Solar Terrestrial and Space Physics, S-RAMP Workshop
Freemantle Western Australia
September 27 – October 2 1998, and

(ii) the South Pacific S-RAMP Meeting
La Trobe University
Bundoora Victoria
28 – 30 September 1999.

All papers were reviewed by at least two independent reviewers.

In addition, papers prepared by members of the Atmospheric and Space Physics group of the Australian Antarctic Division and by members of ANARE, have been added.

This collection of papers is a snapshot of the current direction of Australian solar terrestrial and space physics research in Australia and Antarctica.

The editors thank the respective authors and referees for their valuable contribution, Judy Whelan for her contribution with the preparation of the manuscript text and diagrams, and Cathy Bruce for arranging the final publication.

The papers presented in this edition are available on the Australian Antarctic Division website: www.aad.gov.au and IPS Radio and Space Services website: www.ips.gov.au/STSP.

TABLE OF CONTENTS

1. DISTURBANCES OF THE VERTICAL ELECTRIC FIELD IN HIGH LATITUDES: A REVIEW OF RUSSIAN STUDIES

A. Frank-Kamenetsky and O. Troshichev

Arctic and Antarctic Research Institute
St. Petersburg 199397
Russia

Abstract

A lot of measurements of the vertical geoelectric field have been carried out in the Russian Arctic and in Antarctica during the 1980's. In this paper we briefly review the results of these measurements, supporting the concept of a global electric circuit. These results are not well known to western scientists.

1.1 Introduction

The problem of disturbances in the vertical geoelectric field has been studied for a long time. At first the variations of the potential of the atmospheric electric field at the Earth surface (E_z) were related only to the state of the troposphere, e.g. to the existence of clouds, fogs, rain, industrial aerosols or thunderstorm activity in the locality of the observation point. Later, associations of E_z variations with different heliogeophysical phenomena, such as; solar flares (Reiter, 1971), solar wind density (Sao, 1967), cosmic ray intensity (Reiter, 1972), geomagnetic activity (Rao, 1970) and auroral intensity (Treier, 1961; Olson, 1971; Lobodin and Paramonov, 1972) were found. However, the definite statistical characteristics of these associations have not yet been determined due to the complexity of factors affecting E_z variations, and differences in the methods of measurement and of processing experimental data. The concept of a global electric circuit (Markson, 1978; Roble and Hays, 1978) stimulated a new approach to the problem. According to this theory, tropical thunderstorms are considered as the main electric field generators in the lower atmosphere and they are electrically connected to the ionosphere, and then to the magnetosphere along geomagnetic field lines. High-latitude ionospheric currents are considered to be part of this global circuit, their intensity being affected by the magnitude of the polar cap voltage and the ionospheric conductivity.

1.2 Characteristics of E_z in the auroral zone

1.2.1 Criteria for data selection: 'good weather' conditions

Variations of E_z, caused by local meteorological phenomena, can disguise the large-scale variations due to ionospheric and magnetospheric

influence on the global electric circuit. Analysis carried out by Danilchenko *et al.* (1978) using data from Amderma station (69.9°N, 61.4°E), located in the auroral zone, has shown that the value of E_z decisively depends on the wind velocity, v, and temperature, T. When temperatures are greater than -15°C, E_z can grow to 2.5 kV m^{-1} when the wind velocity increases from zero to 20 m s^{-1}. However when T < -15°C, E_z can reach 20 kV m^{-1} when the wind velocity exceeds 10 m s^{-1}. This effect is evidently related to the growth of the ice crystal concentration in air, which becomes significant when the temperature falls below -12°C. Therefore, the E_z field in the auroral zone turns out to be a minimum in conditions of low wind velocities (< 10 m s^{-1}) and temperatures greater than -15°C. These conditions are taken as the criteria of 'good weather' in analyses of the vertical electric field characteristics in the northern auroral zone.

1.2.2 Latitudinal variations of E_z

Latitudinal variation of E_z and its dependence on magnetic activity has been studied by Bandilet *et al.* (1982a, 1982b) and Zhdanov *et al.* (1984) using data from auroral, sub-auroral and middle latitude stations. Since the magnitude of E_z depends to a large extent on local conditions at each station, the daily-averaged values of E_z (= E_z^D) have been normalised by the quiet level field (= E_z^Q) determined for each station by averaging E_z for 20 days with 'good weather' conditions. Results, presented in Figure 1 for magnetically quiet (ΣK_p < 21, curve 1) and disturbed (ΣK_p > 33, curve 2) days, show that maximal variations,

$$<\delta E_z^D> = <(E_z^D - E_z^Q)/E_z^Q>,$$

are observed at latitudes ϕ > 60°, i.e. in the auroral zone. In magnetically quiet conditions (Figure 1, curve 1) the value $<\delta E_z^D>$ varies rapidly

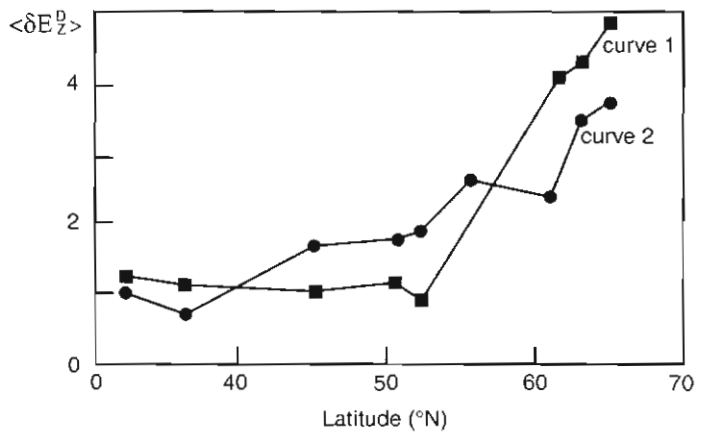


Figure 1. *Latitudinal variation of $<\delta E_z^D>$, for magnetically quiet K_p < 21 (■ curve 1) and disturbed K_p > 33 (● curve 2) conditions in the auroral zone (Bandilet et al., 1982b).*

between latitudes $50° < \phi < 60°$, and deviations of $<\delta E_Z^D>$ are not visible at latitudes less than 50 degrees. Under storm conditions (Figure 1, curve 2), the latitudinal variation of $<\delta E_Z^D>$ is more gradual, and deviations exceeding the quiet level by a factor of two can be observed even at a latitude as low as 45 degrees.

1.2.3 The diurnal variation of E_Z in the auroral zone and its relation to the auroral electrojet

A distinct dawn-dusk asymmetry in diurnal variations of E_Z is typical for stations in the auroral zone (Zhdanov *et al.*, 1984; Pumpyan *et al.*, 1987). As an example, Figure 2 shows the mean diurnal variation of ΔE_Z^H (curve 1), obtained from averaged, hourly measurements from four stations on the Kola peninsula: Lovozero (68.0°N, 35.0°E), Loparskaya (68.9°N, 33.0°E), Apatity (67.7°N, 33.3°E) and Verchnetulomskaya (68.7°N, 31.8°E) using measurements taken in June–July, 1983. The values denoted $<\Delta E_Z^H>$ in Figure 2 were calculated as averaged differences between E_Z^H (= the hourly value of E_Z for a particular day and station) and the daily averaged value (= E_Z^D, = E_Z^H averaged over the 24 hours for the day), $\Delta E_Z^H = (E_Z^H - E_Z^D)/E_Z^D$, and expressed as a percentage.

Only data corresponding to good weather conditions were used in this analysis. According to Pumpyan *et al.* (1987) the diurnal variation in the mean value of ΔE_Z^H (curve 1 in Figure 2) is similar to the diurnal variation of the ionospheric potential, V, calculated by the model of Levitin *et al.* (1984) (curve 2 in Figure 2).

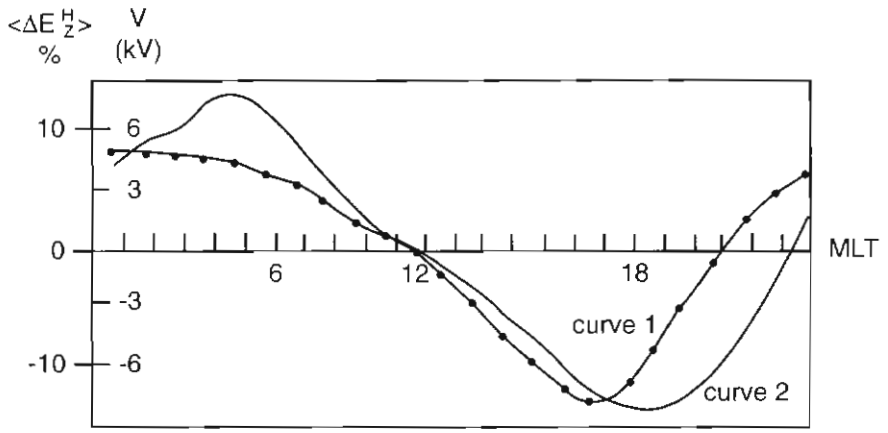


Figure 2. *Diurnal variation of $<\Delta E_Z^H>$, expressed as a percentage, in the auroral zone (• curve 1) and the ionospheric potential (—curve 2) (Pumpyan et al., 1987).*

The dawn-dusk asymmetry of the diurnal variations in mean value of ΔE_Z^H (= $<\Delta E_Z^H>$) grows at all latitudes with an increase in magnetic activity: the diurnal range in $<\Delta E_Z^H>$, expressed as a percentage, is 20–30% for $K_p > 18$ whereas for $K_p > 27$, $<\Delta E_Z^H>$, expressed as a percentage, exceeds 60%. The largest E_Z disturbances (= ΔE_Z^H) are observed within the boundaries of the eastward (1400–2000 UT) and westward (2200–0400 UT) auroral electrojets, and values of E_Z are also maximal here (Bandilet *et al.*, 1982b). This means that maximal E_Z disturbances occur in the area of intense field-aligned currents flowing into and out of the ionosphere (Iijima and Potemra, 1976).

These currents are responsible for the voltage difference across the polar cap. According to the computations of Park (1976a) this potential difference should increase E_Z on the dawn side of the polar cap and decrease E_Z on the dusk side. Therefore experimental data presented in Figure 2 make it possible to conclude that a specific spatial-temporal variation of E_Z in the auroral zone may be linked to magnetosphere-ionosphere processes affecting the global electric circuit. A comparison of E_Z with the auroral electrojets has shown that disturbances in E_Z start during the growth phase of a substorm, 30–40 minutes prior to the beginning of the electrojet equatorward displacement, and maximum values of E_Z are observed one hour before the intense precipitation of auroral particles (Bandilet *et al.*, 1982b).

1.3 Measurements of E_Z in the near pole region

1.3.1 Vostok measurements

Measurements of E_Z in the southern near pole region were carried out at the Russian station Vostok (78.45°S, 106.5°E) between 1974 and 1980 in the framework of a Russian-USA cooperative programme. Vostok station is located in the centre of the polar cap (GML = 84°S), where disturbances in E_Z caused by magnetosphere-ionosphere processes would be the most well defined. A balanced dipole antenna rotating in the magnetic meridian plane was used to measure the electric potential gradient one meter above the snow surface. The first results of these experiments were published by Park (1976b) and showed the influence of the sector structure of the Interplanetary Magnetic Field (IMF) on E_Z variations at Vostok. The study of E_Z characteristics at Vostok station was continued in papers by Frank-Kamenetsky (1983) and Frank-Kamenetsky and Kalita (1986).

1.3.2 'Good weather' conditions at Vostok station

Extensive observations at Vostok have shown that the wind velocity here affects E_Z much more significantly than in the Arctic due to the very low temperatures on the Antarctic Dome. Fluctuations in E_Z, can occur when the wind exceeds the level of about 5 m s^{-1}, (Frank-Kamenetsky, 1983). According to analysis by Frank-Kamenetsky (1983), the criteria for 'good weather' at Vostok are:

(a) absence of snow precipitation,

(b) cloudless sky,

(c) hourly averaged wind speed not significantly different from daily-averaged wind speed,

$$|v^H - v^D| < 0.5 \text{ m s}^{-1},$$

(d) daily-averaged wind speed not much different from monthly-averaged wind speed,

$$|v^D - v^M| < 0.5 \text{ m s}^{-1},$$

(e) hourly-averaged value of E_Z (= E_Z^H) can differ from monthly-averaged value of E_Z for the same hour (= E_Z^M) by not more than E_Z^M, $|E_Z^H - E_Z^M| < E_Z^M$.

1.3.3 Diurnal and seasonal variations

The quiet electric field, E_{Z0}, in days with 'good weather' has been used as a level of reference. This level varies with MLT and season. Maximum values of E_{Z0} were observed in October/November (250 V m^{-1}) and in March/April (150 V m^{-1}). Minimal E_{Z0} values were observed in December/January and June/July (70–120 V m^{-1}). The amplitude of the diurnal variations depends on the season and magnetic activity. In spring (September/October) the maximum values of E_{Z0} during periods of high magnetic activity were observed in the evening hours. The difference between diurnal variations on magnetic disturbed and magnetic quiet days,

$$\delta E_Z = E_{Z0}^{dist} - E_{Z0}^{qui},$$

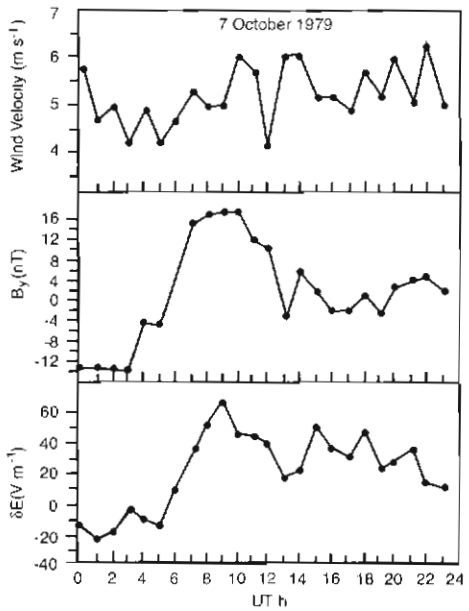


Figure 3. *Example of E_Z disturbance at Vostok station in connection with change in polarity of IMF on 7 October 1979 (Frank-Kamenetsky and Kalita, 1986).*

showed a distinct diurnal variation with an increase of the electric field in the evening hours and a decrease in the morning hours.

1.3.4 Effects of the IMF azimuthal component

Figure 3 gives an example of the influence of the IMF B_y component on the vertical electric field during a period when the ground level wind velocity was stable (V = 5–6 m s^{-1}), on 7 October 1979. δE_z here is the deviation of the electric field from the quiet electric field value, E_{z0}. One can see that δE_z changes from negative to positive values following the turn in polarity of B_y. Figure 4 shows relationship between δE_z and B_y in the day-time period at Vostok station (0800–1600 UT) for some days in October 1979 during which distinct IMF B_y variations occurred. An approximately linear dependence between δE_z and B_y is given by,

$$\delta E_z = (3.1 +/- 0.3)B_y$$

with the correlation coefficient being equal to 0.84 (Frank-Kamenetsky and Kalita, 1986). The conclusion made is that the electric field in the near pole region is strongly affected by the IMF azimuthal component.

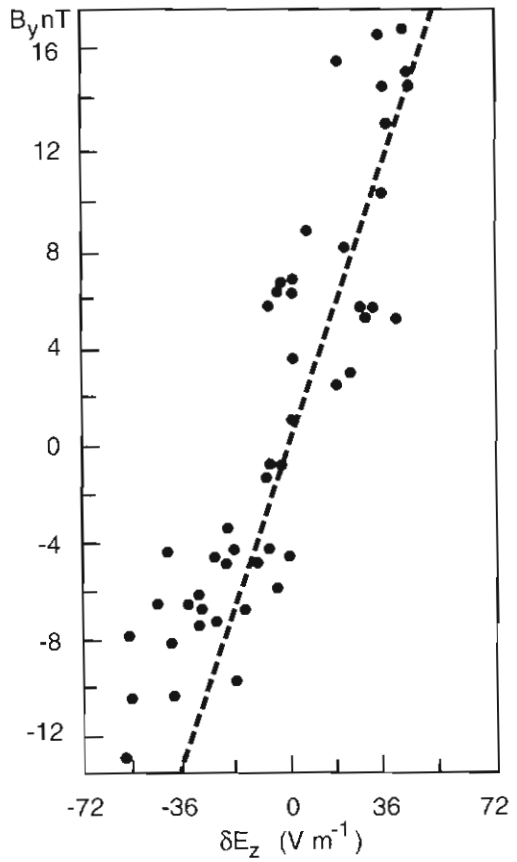


Figure 4. *Relationship between δE_z at Vostok and the IMF B_y component in the day-time period (Frank-Kamenetsky and Kalita, 1986).*

There are some significant differences in the results presented by Park (1976b) and Frank-Kamenetsky and Kalita (1986). The first difference is concerning the seasonal variation of E_z. Park (1976b) indicates that E_z is much stronger in the winter compared to the equinox, whereas Frank-Kamenetsky and Kalita (1986) note that values of E_z are maximum in equinox. The second difference is concerning the influence of IMF polarity. According to Park (1976b) the same effect, a decrease of electric field E_z, is observed when crossing both sector boundaries: toward-to-away ($-B_y$ to $+B_y$) and away-to-toward ($+B_y$ to $-B_y$). Frank-Kamenetsky and Kalita (1986) obtained an opposite response of E_z to different polarities of B_y, i.e. E_z decreases for negative B_y and increases for positive B_y (see Figure 4).

1.4 Conclusion

The results of studies of the vertical electric field E_z carried out in the Russian Arctic and in Antarctica make it possible to conclude that the vertical electric field in high-latitudes is undoubtedly influenced by processes in the magnetosphere-ionosphere system. However, statistical backgrounds of these studies are insufficient to establish with confidence the main characteristics of the high-latitude electric field. Among these characteristics would be the following:

(a) Distinct identification of 'good weather'conditions.

(b) Determination of diurnal variation of E_z.

(c) Determination of seasonal variation of E_z.

(d) Relationship between E_z and wind velocity at ground level.

(e) Relationship between E_z and IMF parameters (B_y and E_z).

(f) Relationship between E_z and magnetic activity.

An expanded base of digital experimental data would be required for statistically well-founded analyses.

1.5 Acknowledgement

The preparation of this report has been funded by an Australian Antarctic Foundation grant.

References

Bandilet, O.I., Kozelova, T.V., Chernysheva, S.P. and Sheftel, V.M. (1982a). About nature of vertical component of electric field during the magnetic disturbances. *Geomagnetism and Aeronomy,* 22: p. 1012 (in Russian).

Bandilet, O.I., Frank-Kamenetsky, A.V., Chernysheva, S.P. and Sheftel, V.M. (1982b). Space distribution of disturbances of vertical component of geoelectric field accompanying the magnetic activity increase. *Geomagnetism and Aeronomy,* 22. p. 1028 (in Russian).

Danilchenko, D.P., Zemlyankin, G.I., Chernysheva, S.P. and Sheftel, V.M. (1978). Data sampling in analysis of auroral effects in geoelectric field. *Geomagnnetism and Aeronomy,* 18. p. 1121 (in Russian).

Frank-Kamenetsky, A.V. (1983). Features of the vertical atmospheric electric field at Vostok station. In: *Atmospheric electricity and magnetospheric disturbances, Moscow.* p. 91 (in Russian).

Frank-Kamenetsky, A.V. and Kalita, V.M. (1986). The connection between atmospheric electric field variations in the polar cap and IMF parameters. *Proceedings of AARI, Leningrad, Hydrometeoizdat,* 405. p. 48 (in Russian).

Iijima, T. and Potemra, T.A. (1976). Field-aligned currents in the day-side cusp observed by TRIAD. *Journal of Geophysical Research,* 81: 5971–5979.

Levitin, A.E., Feldstein, Ya. I. and Afonina, R.G. (1984). Model of large-scale electric fields and currents in the high-latitude ionosphere. Moscow, Hydrometeoizdat. p. 178 (in Russian).

Lobodin, T.V. and Paramonov, N.A. (1972). Variation of atmospheric electric field during aurora. *Pure and Applied Geophysics,* 100: 167–173.

Markson, R. (1978). Atmospheric electricity and problem of relationship between solar activity and weather. In: McCormac, B. and Seliga, T. (Eds.). *Proceedings of the Symposium on Solar-Terrestrial Influences in Weather and Climate.* p. 215.

Olson, D.E. (1971). The evidence for auroral effects on atmospheric electricity. *Pure and Applied Geophysics,* 84: 118–138.

Park, C.G. (1976a). Downward mapping of high-latitude ionospheric electric fields to the ground. *Journal of Geophysical Research,* 81: 168–174.

Park, C.G. (1976b). Solar magnetic sector effects on the vertical atmospheric electric field at Vostok, Antarctica. *Geophysical Research Letters,* 3: 475–478.

Pumpyan, P.E., Chernysheva, S.P., Sheftel, V.M. and Yaroshenko, A.N. (1987). Solar-magnetospheric effects in variations of the atmospheric electric field in the high-latitude zone. *Geomagnetism and Aeronomy,* 27. p. 287 (in Russian).

Rao, M. (1970). On the possible influence of magnetic activity on the atmospheric electric parameters. *Journal of Atmospheric and Terrestrial Physics,* 32: 1431–1437.

Reiter, R. (1971). Further evidence for impact of solar flares on potential gradient and air-earth current characteristics at high mountains. *Pure and Applied Geophysics,* 86: 142–158.

Reiter, R. (1972). Case study concerning the impact of solar activity upon potential gradient and air-earth current in the lower troposphere. *Pure and Applied Geophysics,* 94: 218–225.

Roble, R.G. and Hays, P.B. (1978). Electric coupling between the upper and lower atmosphere, In: McCormac, B. and Seliga, T. (Eds.). *Proceedings of the Symposium on Solar-Terrestrial Influences on Weather and Climate.* p. 223.

Sao, K. (1967). Correlations between solar activity and the atmospheric potential gradient at the Earth's surface in the polar regions. *Journal of Atmospheric and Terrestrial Physics,* 29: 213–215.

Freier, G.D. (1961). Auroral effects on the Earth's electric field. *Journal of Geophysical Research,* 66: 2695–2702.

Zhdanov, P.F., Moiseev, V.G., Novikov, A.M., Sokolov, V.D., Frank-Kamenetsky, A.V., Chemyshova, S.P., Sheftel, V.M. and Yaroshenko, A.N. (1984). Influence of electric fields of magnetospheric convection on variations of atmospheric electric field in a high-latitude zone. *Geomagnetism and Aeronomy,* 24. p. 844 (in Russian).

2. GLOBAL GRAVITY WAVE 'WEATHER' IN THE MIDDLE ATMOSPHERE: PRELIMINARY INSIGHTS FROM THE CRISTA-SPAS MISSIONS

S.D. Eckermann[1], P. Preusse[2], B. Schaeler[2], J. Oberheide[2], D. Offerman[2], J.T. Bacmeister[2] and D. Broutman[4]

(1) E.O. Hulburt Center for Space Research
Code 7641.2 Naval Research Laboratory
Washington DC 20375 USA
(email: eckerman@map.nrl.navy.mil)

(2) Department of Physics
University of Wuppertal
Gauss-Str. 20
D-42097 Wuppertal Germany

(3) Universities Space Research Association
10227 Wincopin Circle
Suite #202
Columbia MD 21044 USA

(4) Computational Physics Inc.
Suite #600
2750 Prosperity Avenue
Fairfax VA 22031 USA
and
School of Mathematics
University of New South Wales
Sydney NSW 2052 Australia

Abstract

We describe a preliminary analysis of small-scale temperature perturbations in the stratosphere produced by long wavelength gravity waves, as measured globally during the CRISTA-SPAS missions. We focus on enhanced activity observed at the equator and near the southern tip of South America. Global ray-tracing simulations for the mission days indicate that the equatorial activity is broadly consistent with the transmission of nonzero phase speed waves into the stratosphere. Ray-tracing and mountain wave model simulations indicate that the activity over South America is produced by long wavelength mountain waves forced by flow over the Andes. The results suggest that the study of global gravity wave 'weather' in the middle atmosphere may soon be feasible.

2.1 Introduction

A range of weather-related phenomena in the lower atmosphere generates gravity waves. Examples include fronts (e.g. Eckermann and Vincent, 1993), squall lines (e.g. Alexander *et al.*, 1995), convective clouds (e.g. Alexander and Pfister 1995), cyclogenesis (e.g. Powers, 1997), typhoons (e.g. Sato, 1993), geostrophic adjustment of jet steams (e.g. Kaplan *et al.*, 1997), flow over mountains (e.g. Ralph *et al.*, 1993), and so on. If wind patterns aloft are favourable, some of these waves can propagate into the middle atmosphere. Due to decreasing atmospheric densities, gravity waves grow in amplitude with height and so must eventually generate instabilities and 'break.' Such quasi-continuous gravity wave breaking generates drag and diffusion that strongly influence the global circulation and temperature patterns of the middle atmosphere (e.g. Haynes *et al.*, 1991; Fritts and Luo, 1995; Alexander and Holton, 1997).

While the impact of breaking gravity waves on the climatology of the middle atmosphere is well appreciated, far less is known about how shorter-term changes in weather-related gravity wave sources and local propagation environments affect the 'weather' in the middle atmosphere. Such research has been stymied to date by a lack of data on the synoptic meteorology of the middle atmosphere and its contemporaneous gravity-wave activity. There have also been very few global models of gravity-wave production and propagation with any sort of forecasting capability.

These shortcomings in observations and modelling are slowly being resolved, allowing us to investigate 'gravity wave weather' issues systematically for the first time. Next generation satellite instruments have sufficient spatial resolution to resolve explicitly large-scale gravity waves in the middle atmosphere (e.g. Ross *et al.*, 1992; Fetzer and Gille, 1994; Mende *et al.*, 1994; Wu and Waters, 1996; Dewan *et al.*, 1998; Preusse *et al.*, 1999; Eckermann and Preusse, 1999). Similar advances in modelling have led to global models that can simulate and even forecast gravity wave activity generated by specific sources (e.g. Bacmeister *et al.*, 1994; Marks and Eckermann, 1995; Alexander and Holton, 1997; Dörnbrack *et al.*, 1998; Sato *et al.*, 1999; Charron and Brunet, 1999). This has enabled some initial comparisons between global models and global measurements of gravity waves in the middle atmosphere (Eckermann and Marks, 1997; Alexander, 1998; Eckermann and Preusse, 1999).

We focus here on preliminary data from two observational campaigns with the CRISTA-SPAS satellite during shuttle missions STS-66 and STS-85 (November 1994 and August 1997, respectively). Analysis shows that large-scale gravity wave fluctuations are present in the stratospheric and mesospheric temperature measurements. These conclusions are buttressed by numerical models, which generate maps of gravity wave activity that resemble the measured fluctuations. These results provide some preliminary insights into the generation and filtering processes affecting gravity waves in the middle atmosphere.

2.2 The CRISTA-SPAS missions

The Cryogenic Infrared Spectrometers and Telescopes for the Atmosphere (CRISTA) instrument operated aboard the Shuttle Pallet Satellite (SPAS) (Offermann *et al.*, 1999). CRISTA-SPAS was deployed into orbit and later recovered by the shuttle during missions STS-66 (3–14 November 1994) and STS-85 (7–16 August 1997). Figure 1 shows deployment of CRISTA-SPAS by the orbiter's robot arm on 4 November 1994. Lee wave clouds downstream (west) of the Andean ridge over the Chubut River region in central Patagonia (Argentina) are visible in this photograph, indicating that mountain waves were being generated by flow over the Andes as CRISTA-SPAS was being prepared for orbit.

Figure 1. *Astronaut photo of CRISTA-SPAS being deployed into orbit by the robot arm of the Space Shuttle Atlantis at 1219 Z on 4 November 1994. Visible below are lee wave clouds downstream of the southern Andes over the Chubut region of Argentina (~ 42°S). This photograph and others from STS-66 are accessible from NASA's shuttle image archive, at: http://images.jsc.nasa.gov/iams/images/earth/STS066/html/20177512.htm.*

CRISTA is a limb-scanning instrument that measures infrared emission spectra with high spectral resolution (Riese *et al.*, 1999; Grossmann, 2000). Temperatures in the stratosphere and lower mesosphere are retrieved from the CO_2 Q-branch emission at 792 cm^{-1} (~ 12.6 μm). Full details of the 'version 1' retrievals are given by

Riese *et al.* (1999), which yielded temperatures with a precision ~ 1 K. The precision of the 'version 2' data analysed here is ~ 0.5 K. CRISTA uses two additional telescopes that look 18° either side of the central telescope, corresponding to tangent heights separated by ~ 650 km. Each scanning telescope acquires a complete vertical profile every 200–400 km along track, depending on the measuring mode. These unique measurements yielded data with global coverage and high spatial resolution (see Figure 2 of Riese *et al.*, 1999). Mesospheric temperatures are retrieved from CO_2 emissions in the 652–657 cm^{-1} band, which are measured by the centre telescope only. Their precision is estimated to be ~ 1 K. The vertical sampling interval is 1.5 km for both the stratospheric and mesospheric temperature data. Further details of the instrument and measurements are given by Offermann *et al.* (1999), Riese *et al.* (1999) and Grossmann (2000).

Fetzer and Gille (1994, 1996) found large-scale gravity-wave fluctuations in stratospheric data acquired by the Limb Infrared Monitor of the Stratosphere (LIMS), which operated on the Nimbus 7 satellite during 1978–79. The precision and resolution of CRISTA measurements exceed those of LIMS. Furthermore, CRISTA measures spectra whereas LIMS was a filter instrument. This allows CRISTA to check the quality of the retrievals in regions of enhanced aerosol content (e.g. the tropical lower stratosphere). Retrieval modelling by Preusse *et al.* (1999) indicates that sinusoidal gravity wave temperature perturbations with wavelengths in excess of ~ 5 km vertically and ~ 200 km horizontally are resolved by the CRISTA temperature channels. Thus, gravity waves with wavelengths longer than these limits should be evident in the CRISTA temperature data.

To investigate this, we isolated small-scale temperature variability in the CRISTA temperature data. Synoptic variability was removed using a wavenumber 0–6 Kalman filter. Next, vertical profiles of residual fluctuations were analysed using the Maximum Entropy Method (MEM) and harmonic analysis. MEM analysis was used to identify the energetic vertical wavelength oscillations in a given profile. These wavelengths were then used to constrain a harmonic fit over a 12 km altitude interval, which was progressively shifted over the full height range of the profile. In this way, two principal amplitudes and phases were determined as a function of height for each profile. Finally, a multiplier, determined by the instrument's 'visibility' to this particular vertical wavelength, was used to scale the fitted amplitude. For further details, see Preusse *et al.* (1999) and Eckermann and Preusse (1999).

Figure 2 shows resulting global maps of small-scale temperature amplitudes at 20–30 km for six successive days during Mission 1 (STS-66). We note a band of enhanced variance at the equator that exhibits significant zonal asymmetry and day-to-day variability. This is superficially consistent with the notion of enhanced convective launching of gravity waves and equatorial waves into the tropical stratosphere (e.g. Alexander and Pfister, 1995; Garcia, 2000). We also note a

persistent zone of enhanced activity near the southern tip of South America, consistent with the concept of mountain waves generated by flow over the southern Andes (see Figure 1) penetrating into the stratosphere. However, work by Alexander (1998) has demonstrated that straightforward interpretations of raw satellite gravity wave data such as these can be potentially misleading. Therefore we turn to global numerical modelling of gravity waves to make further progress in interpreting these and other features.

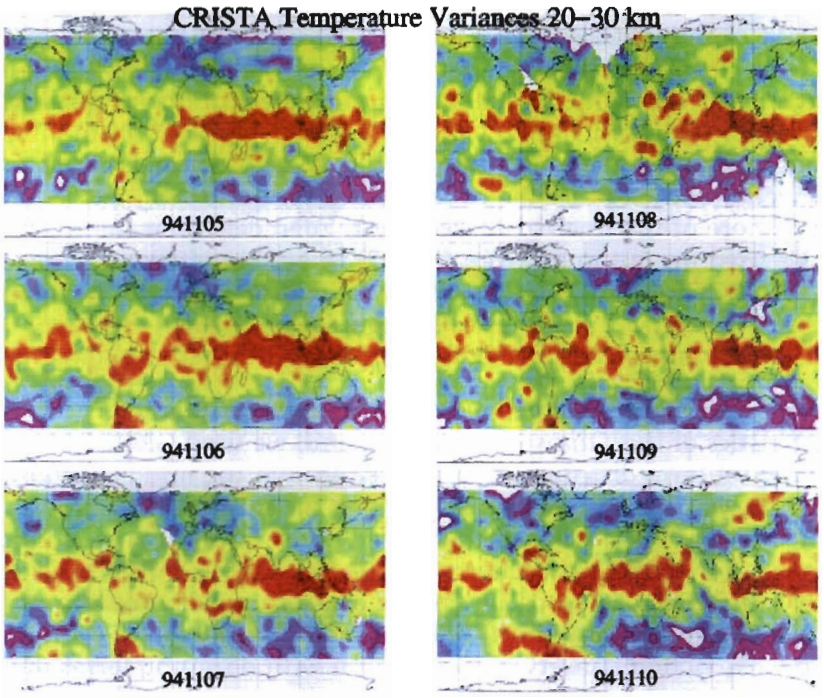


Figure 2. *Preliminary global maps of CRISTA temperature amplitudes in the 20–30 km height range for six successive days during Mission 1 (5–10 November 1994). Red (dark) shading indicates highest values (~ 2.5 K), and blue-white shading shows smallest values.*

2.3 Global gravity wave models

2.3.1 Gravity Wave Regional or Global Ray Tracer (GROGRAT)

GROGRAT is a numerical model that uses ray-tracing techniques to track the propagation and amplitude evolution of nonhydrostatic gravity waves within arbitrary regional or global specifications of the Earth's atmosphere. Marks and Eckermann (1995) described the basic formulation of the model, while Eckermann and Marks

(1997) reported subsequent improvements and updates to the code. Further information can be found at the GROGRAT web site: http://uap-www.nrl.navy.mil/dynamics/html/grograt.html.

For the background atmosphere, we use 6 hourly global 2.5° x 2° (144 x 91) wind and temperature data from NASA's Data Assimilation Office (DAO), which come gridded at 18 standard pressure levels from 1000 to 0.4 hPa (Schubert *et al.*, 1993; Coy and Swinbank, 1997). We resampled these data onto a 64 x 45 x 18 global grid. Comparison of Figures 3a and 3b shows that the major weather phenomena are retained in the wind fields after this resampling. GROGRAT absorbed these fields, first regridding them onto a geometrical height grid (using the DAO geopotential height fields), then fitting the data at each height with a superposition of spherical harmonics $Ym, n(\varphi,\theta)$. Lower panels in Figure 3 show fits using spherical harmonics out to orders (m, n) = (16, 16) and (m, n) = (32, 32) (panels c and d, respectively). Both capture the overall structure quite well, but Figure 3d captures more of the finer-scale weather-related details in the winds. Thus, we used fits out to (m, n) = (32, 32) in our simulations. Note also that GROGRAT interpolates all the individual spherical harmonic coefficients at each fitting level with cubic splines, which ensures smoothly varying continuous derivatives in all three spatial directions.

Here we ignore sources and concentrate on the transmission characteristics of the atmosphere. To this end, we used GROGRAT to trace a globally uniform distribution of waves through the middle atmosphere using DAO winds and temperatures during the mission days. Waves were launched at z_0 = 1 km with ground-based phase speeds c = 0, 20 and 40 m s^{-1}, a horizontal wavelength λ_x = 50 km, a peak initial horizontal velocity amplitude of u' = 0.2 m s^{-1} and an isotropic distribution of eight initial propagation directions. Wave amplitudes were controlled using a wave action conservation equation with damping terms due to turbulent diffusion, infrared radiative cooling and wave breaking/saturation (Marks and Eckermann, 1995). A total of 13056 rays were traced over the globe.

Figure 4 shows global ray counts, weighted by the horizontal velocity amplitude of each ray, as a function of height on 6 November 1994 at 0000 UT. Results shown are for the c = 20 m s^{-1} waves, as might plausibly emanate from convection and other moving and/or evolving weather systems. Although wave amplitudes are globally uniform at z_0 = 1 km, the variable background atmosphere quickly yields considerable geographical variability at upper heights. At z = 20 km and above, Figure 4 reveals a clear band of enhanced activity around the equator that resembles the enhanced equatorial activity at 20–30 km in the CRISTA data (Figure 2). It occurs in the model due to very light winds throughout the equatorial troposphere and stratosphere during Mission 1, which enables waves with nonzero phase speeds to propagate upwards without encountering critical layers or severe damping. However, the zonal asymmetry in Figure 2 is not reproduced in

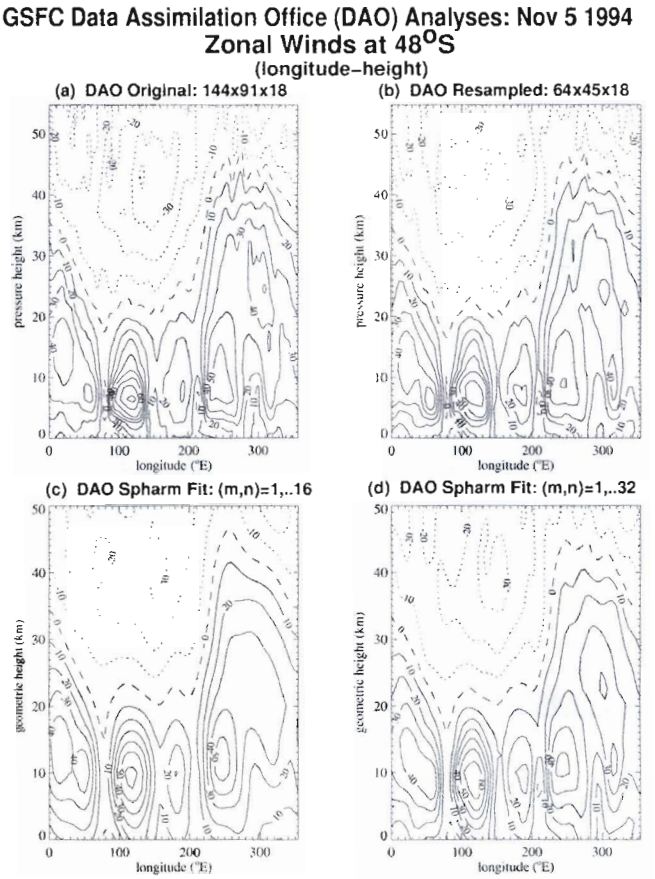


Figure 3. *Longitude-height plot of DAO zonal wind at 48°S on 5 November 1994 at 0000 Z: (a) original data; (b) resampled to 64 x 45 in the horizontal; (c) after GROGRAT regridding to geometric heights and spherical harmonic fits to maximum (m, n) = (16, 16); (d) as for (c), but after spherical harmonic fits to maximum (m, n) = (32, 32).*

Figure 4, suggesting that zonally varying wave generation may also play a role in producing the observed activity.

We are investigating possible sources of equatorial wave activity during the missions. For example, during Mission 2 (STS-85; 7–16 August 1997) similar bands of enhanced equatorial activity were observed. Super Typhoon Winnie developed over the South China Sea during the mission, reaching Category 5 ('super typhoon' status) on August 12 (see Figure 5). Typhoons are known to be strong sources of gravity waves (e.g. Matsumoto and Okamura, 1985; Sato, 1993), and thus Winnie might have produced enhanced stratospheric gravity wave activity. CRISTA used a 'hawkeye' observational mode during Mission 2, which increased sampling over Indonesia but reduced overall coverage over the typhoon. Nevertheless, preliminary observations in this region

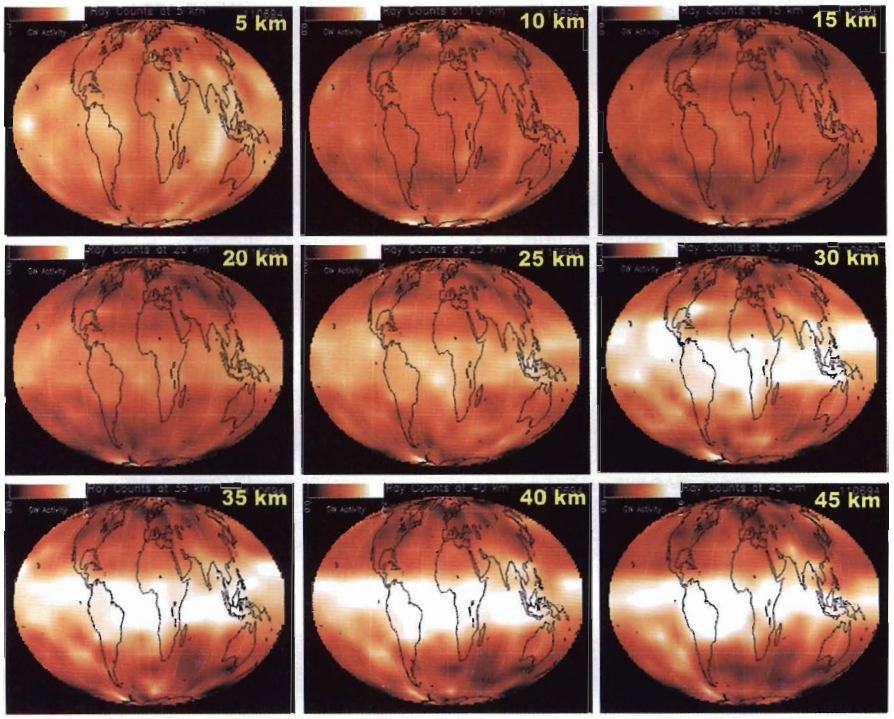


Figure 4. *GROGRAT simulations of global gravity wave activity on 6 November 1994, using a globally invariant source spectrum of waves of c = 20 m s^{-1} (see text). Light values indicate large wave variances, darker values indicate small wave variances.*

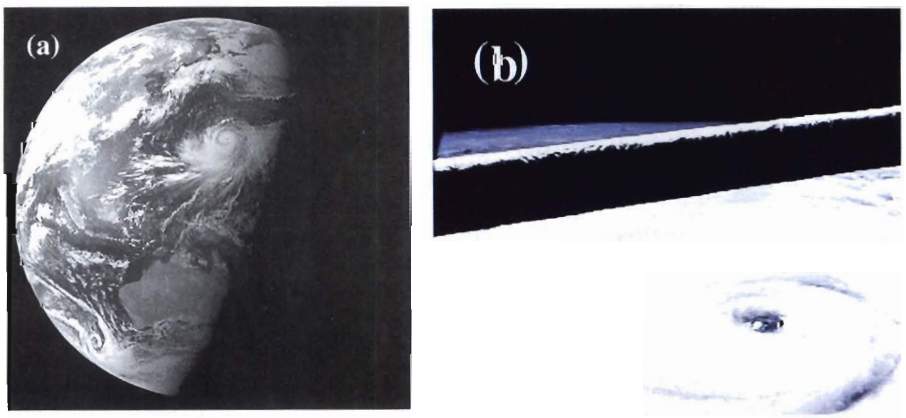


Figure 5. *(a) Full disk GOES 8 visible image on 12 August 1997, showing Super Typhoon Winnie. Australia is visible to the south (image courtesy of NOAA's National Climatic Data Center), and (b) close-up photo of Winnie from Discovery on 13 August (robot arm in the foreground).*

show some suggestions of enhanced temperature fluctuations in the stratosphere near the typhoon. Investigation of equatorial source processes continues.

2.3.2 Naval Research Laboratory Mountain Wave Forecast Model (MWFM)

Mountains are believed to be a strong extratropical source of gravity waves for the middle atmosphere (Nastrom and Fritts, 1992; Bacmeister, 1993). With this in mind, the Naval Research Laboratory Mountain Wave Forecast Model (MWFM) was developed to forecast the turbulence produced by breaking mountain waves in the stratosphere, as an aid to safe flight planning for NASA's stratospheric ER-2 research aircraft (Bacmeister *et al.*, 1994; Eckermann *et al.*, 2000). The model uses a database of quasi-two-dimensional ridges inferred from high-resolution digital topography. Surface winds from forecast data are blown across these ridges and used to generate waves. Wave equations are used thereafter to trace any vertical extension of this activity into the atmosphere, as well as any subsequent wave breaking. Further information can be found at the MWFM web site: http://uap-www.nrl.navy.mil/dynamics/html/mwfm.html.

Here we use the MWFM with DAO data to generate an initial global field of mountain waves forced at the ground. We use GROGRAT thereafter to track the propagation of these waves to greater heights, using fitted DAO data as before. Figure 6 shows modelled mountain wave activity at z = 25 km from 5–10 November 1994. We note that stratospheric mountain wave activity is predicted at the southern tip of South America during Mission 1, which resembles the activity observed by CRISTA in Figure 2. However, there is additional activity over Eurasia and western North America in Figure 6 that is not seen as clearly in Figure 2.

In analysing MLS radiance fluctuations, Alexander (1998) stressed the importance of accounting for the spatial filtering of the satellite gravity wave measurement. Mountain wave vertical wavelengths are given approximately by

$$\lambda_z = \frac{2\pi U \cos\varphi}{N} \qquad (1)$$

where U is the horizontal wind speed, N is the Brunt-Väisälä frequency and φ is the angle between the horizontal wind and wave vectors. Between 40°–50°S over South America (~ 70°W) the DAO winds blow approximately orthogonal to the Andean ridge axis, implying $\varphi \sim 0°$. At 20–30 km, U ~ 20–50 m s^{-1} (see Figure 3) and N ~ 0.02 rad s^{-1}, yielding λ_z ~ 6–16 km from Equation 1. Conversely, U ~ 10 m s^{-1} at 30°N, yielding vertical wavelengths no greater than ~ 3 km here. Since CRISTA is sensitive to vertical wavelengths $\lambda_z > 5$ km, it can resolve the long λ_z waves over South America but not the shorter λ_z waves at northern latitudes.

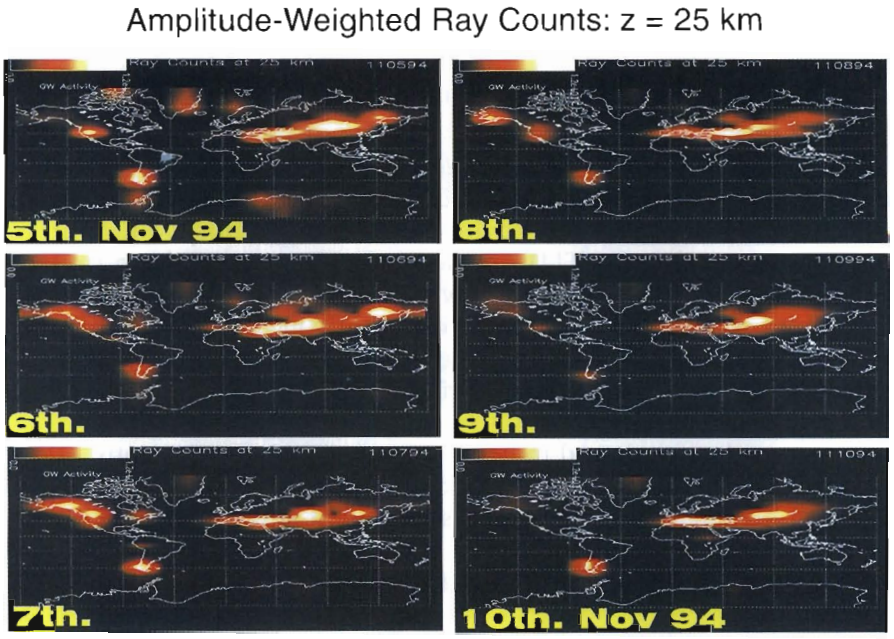


Figure 6. *Coupled GROGRAT-NRL/MWFM simulations of global mountain wave amplitudes at z = 25 km, using DAO winds and temperatures. Light (dark) values depict large (small) wave activity.*

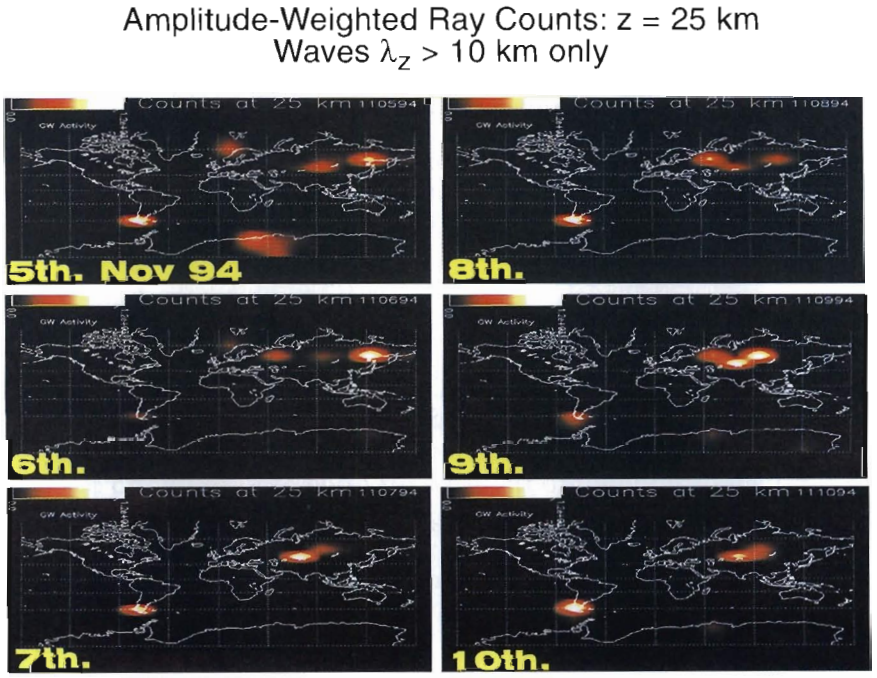


Figure 7. *As in Figure 6, but after imposition of a filter which retains λz > 10 km waves only.*

To assess this quantitatively, Figure 7 shows the model results from Figure 6 after imposition of a filter which retains only those waves with vertical wavelengths $\lambda_z > 10$ km. We see that the activity over southern South America is now more prominent and activity in the northern hemisphere is significantly attenuated, which is more in line with the data in Figure 2. Alexander (1998) reported similarly improvements on comparing filtered model data to the MLS radiance fluctuations reported by Wu and Waters (1996).

Eckermann and Preusse (1999) present an in-depth study of temperature fluctuations observed by CRISTA over southern South America and central Eurasia. Using simple mountain wave theory and 'next generation' MWFM hindcast simulations, they show conclusively that this activity is produced by long wavelength mountain waves which propagate into the stratosphere. Mesoscale modelling work by Tan and Eckermann (2000) generates stratospheric mountain waves very similar to those measured by CRISTA over the southern Andes on 6 November 1994.

2.4 Summary

Initial comparisons between the small-scale temperature perturbations measured by CRISTA (Figure 2) and global gravity wave model predictions are encouragingly favourable. These findings further support the conclusion that CRISTA temperature measurements in the stratosphere and mesosphere resolved large-scale gravity waves (see also Preusse *et al.*, 1999; Eckermann and Preusse, 1999). Global synoptic data on gravity waves in the middle atmosphere are still very limited, having been inferred from a few earlier satellite measurements only (Fetzer and Gille, 1994, 1996; Wu and Waters, 1996). Furthermore, interpretation of some of these earlier measurements has not always been straightforward (Alexander, 1998). Yet global measurements of gravity waves and accompanying theoretical explanations of the observed features are vital to a proper understanding of the 'weather' in the middle atmosphere, which is very sensitive to gravity wave driving from below. Our analysis suggests that some key features in the CRISTA data are consistent with specific gravity wave production and filtering processes. Further analysis and modelling of these data are being pursued.

2.5 Acknowledgements

This research was supported in part by the Office of Naval Research and by NASA through grants NAS5-98045 and L68786D. Thanks to Iain Reid for the invitation and support to attend the AIP conference and the University of Adelaide.

References

Alexander, M.J. (1998). Interpretations of observed climatological patterns in stratospheric gravity wave variance. *Journal of Geophysical Research*, 103: 8627–8640.

Alexander, M.J. and Holton, J.R. (1997). A model study of the zonal forcing in the equatorial stratosphere by convectively induced gravity waves. *Journal of the Atmospheric Sciences*, 54: 1408–419.

Alexander, M.J. and Pfister, L. (1995). Gravity wave momentum flux in the lower stratosphere over convection. *Geophysical Research Letters*, 22: 2029–2032.

Alexander, M.J., Holton, J.R. and Durran, D.R. (1995). The gravity wave response above deep convection in a squall line simulation. *Journal of the Atmospheric Sciences*, 52: 2212–2226.

Bacmeister, J.T. (1993). Mountain-wave drag in the stratosphere and mesosphere inferred from observed winds and a simple mountain-wave parameterization scheme. *Journal of the Atmospheric Sciences*, 50: 377–399.

Bacmeister, J.T., Newman, P.A., Gary, B.L. and Chan, K.R. (1994). An algorithm for forecasting mountain wave-related turbulence in the stratosphere. *Weather and Forecasting*, 9: 241–253.

Charron, M. and Brunet, G. (1999). Gravity wave diagnosis using empirical normal modes. *Journal of the Atmospheric Sciences*, 56: 2706–2727.

Coy, L. and Swinbank, R. (1997). Characteristics of stratospheric winds and temperatures produced by data assimilation. *Journal of Geophysical Research*, 102: 25763–25781.

Dewan, E.M., Picard, R.H., O'Neil, R.R., Gardiner, H.A., Gibson, J., Mill, J.D., Richards, E., Kendra, M. and Gallery, W.O. (1998). MSX satellite observations of thunderstorm-generated gravity waves in mid-wave infrared images of the upper stratosphere. *Geophysical Research Letters*, 25: 939–942.

Dörnbrack, A., Leutbecher, M., Volkert, H. and Wirth, M. (1998). Mesoscale forecasts of stratospheric mountain waves. *Meteorological Applications*, 5: 117–126.

Eckermann, S.D. and Marks, C.J. (1997). GROGRAT: A new model of the global propagation and dissipation of atmospheric gravity waves. *Advances in Space Research*, 20(6): 1253–1256.

Eckermann, S.D. and Preusse, P. (1999). Global measurements of stratospheric mountain waves from space. *Science*, 286: 1534–1537.

Eckermann, S.D. and Vincent, R.A. (1993). VHF radar observations of gravity-wave production by cold fronts over southern Australia. *Journal of the Atmospheric Sciences*, 50: 785–806.

Eckermann, S.D., Broutman, D., Tan, K.A., Preusse, P. and Bacmeister, J.T. (2000). Mountain waves in the stratosphere. *NRL Review* (in press).

Fetzer, E.J. and Gille, J.C. (1994). Gravity wave variance in LIMS temperatures. Part I: variability and comparison with background winds. *Journal of the Atmospheric Sciences*, 51: 2461–2483.

Fetzer, E.J. and Gille, J.C. (1996). Gravity wave variance in LIMS temperatures. Part II: comparison with the zonal-mean momentum balance. *Journal of the Atmospheric Sciences*, 53: 398–410.

Fritts, D.C. and Luo, Z. (1995). Dynamical and radiative forcing of the summer mesopause circulation and thermal structure, 1, mean solstice conditions. *Journal of Geophysical Research*, 100: 3119–3128.

Garcia, R.R. (2000). The role of equatorial waves in the semiannual oscillation of the middle atmosphere. In: Siskind, D.E., Eckermann, S.D. and Summers, M.E. (Eds.). *Atmospheric Science Across the Stratopause*. American Geophysical Union Monograph, V. 123.

Grossmann, K.U. (2000). Recent improvements in middle atmosphere remote sounding techniques: the CRISTA-SPAS experiment. In: Siskind, D.E., Eckermann, S.D. and Summers, M.E. (Eds.). *Atmospheric Science Across the Stratopause*. American Geophysical Union Monograph, V. 123.

Haynes, P.H., Marks, C.J., McIntyre, M.E., Shepherd, T.G. and Shine, K.P. (1991). On the "downward control" of extratropical diabatic circulations by eddy-induced mean forces. *Journal of the Atmospheric Sciences*, 48: 651–678.

Kaplan, M.L., Koch, S.E., Lin, Y.-L., Weglarz, R.P. and Rozumalski, R.A. (1997). Numerical simulations of a gravity wave event over CCOPE, I, the role of geostrophic adjustment in mesoscale jetlet formation. *Monthly Weather Review*, 125: 1185–1211.

Marks, C.J. and Eckermann, S.D. (1995). A three-dimensional nonhydrostatic ray-tracing model for gravity waves: formulation and preliminary results for the middle atmosphere. *Journal of the Atmospheric Sciences*, 52: 1959–1984.

Matsumoto, S. and Okamura, H. (1985). The internal gravity wave observed in the Typhoon T8124 (Gay). *Journal of the Meteorological Society of Japan*, 63: 37–51.

Mende, S.B., Swenson, G.R., Geller, S.P. and Spear, K.A. (1994). Topside observations of gravity waves. *Geophysical Research Letters*, 21: 2283–2286.

Nastrom, G.D. and Fritts, D.C. (1992). Sources of mesoscale variability of gravity waves, 1, topographic excitation. *Journal of the Atmospheric Sciences*, 49: 101–110.

Offermann, D., Grossmann, K.U., Barthol, P., Knieling, P., Riese, M. and Trant, R. (1999). The cryogenic infrared spectrometers and telescopes for the atmosphere (CRISTA) experiment and middle atmosphere variability. *Journal of Geophysical Research*, 104: 16311–16325.

Powers, J.G. (1997). Numerical model simulations of a mesoscale gravity wave event: sensitivity tests and spectral analysis. *Monthly Weather Review*, 125: 1838–1869.

Preusse, P., Schaeler, B., Bacmeister, J.T. and Offermann, D. (1999). Evidence for gravity waves in CRISTA temperatures. *Advances in Space Research*, 24: 1601–1604.

Ralph, F.M., Neiman, P.J. and Levinson, D. (1993). Lidar observations of a breaking mountain waves associated with extreme turbulence. *Geophysical Research Letters*, 24: 663–666.

Riese, M., Spang, R., Preusse, P., Ern, M., Jarisch, M., Offerman, D. and Grossman, K.U. (1999). Cryogenic infrared spectrometers and telescopes for the atmosphere (CRISTA) data processing and atmospheric temperature and trace gas retrieval. *Journal of Geophysical Research*, 104: 16349–16367.

Ross, M.N., Christensen, A.B., Meng, C.-I. and Carbary, J.F. (1992). Structure in the UV nightglow observed from low Earth orbit. *Geophysical Research Letters*, 19: 985–988.

Sato, K. (1993). Small-scale wind disturbances observed by the MU radar during the passage of Typhoon Kelly. *Journal of the Atmospheric Sciences*, 50: 518–537.

Sato, K., Kumakura, T. and Takahashi, M. (1999). Gravity waves appearing in a high-resolution GCM simulation. *Journal of the Atmospheric Sciences*, 56: 1005–1018.

Schubert, S.D., Rood, R.B. and Pfaendtner, J. (1993). An assimilated dataset for Earth science applications. *Bulletin of the American Meteorological Society*, 74: 2331–2342.

Tan, K.A. and Eckermann, S.D. (2000). Numerical model simulations of mountain waves in the middle atmosphere over the southern Andes. In: Siskind, D.E., Eckermann, S.D. and Summers, M.E. (Eds.). *Atmospheric Science Across the Stratopause*. American Geophysical Union Monograph, V.123.

Wu, D.L. and Waters, J.W. (1996). Gravity-wave-scale temperature fluctuations seen by the UARS MLS. *Geophysical Research Letters*, 23: 3289–3292.

3. RAYLEIGH LIDAR MEASUREMENTS OF MEAN STRATOPAUSE REGION TEMPERATURES ABOVE KINGSTON, TASMANIA (43.0°S, 147.3°E) BETWEEN MARCH 1997 AND APRIL 1999

A.R. Klekociuk

Atmospheric and Space Physics
Australian Antarctic Division
Kingston Tasmania 7050 Australia
(email: andrew.klekociuk@aad.gov.au)

Abstract

A Rayleigh lidar was operated at Kingston, Tasmania on 22 nights between March 1997 and April 1999. During this interval, profiles of scattering ratio obtained from the lidar data indicate that the stratospheric aerosol layer extended up to a maximum altitude of about 34 km. The standard Rayleigh lidar inversion technique was applied to the data to yield profiles of temperature between the top of the aerosol layer and 60 km altitude. Mean temperatures in the stratopause region show a seasonal trend consistent with predictions of the MSISE-90 climatological model.

3.1 Introduction

Temperature measurements in the middle atmosphere (15 km to 90 km altitude) have been obtained by a variety of *in situ* and remote sensing techniques. The technique that arguably holds the most promise in terms of accuracy and resolution is lidar. Conventionally, lidar soundings of Rayleigh or Raman backscatter are used to derive the density profile of atmospheric air molecules, which in turn yields the temperature profile under the assumption that the atmosphere is in hydrostatic equilibrium and obeys the perfect gas law (Hauchecorne and Chanin, 1980).

There is growing interest in middle atmosphere temperature measurements owing to concerns and uncertainties relating to global climate change. Importantly, the middle atmosphere may be a litmus site for detecting anthropogenic influences on climate. Modelling suggests that cooling of the middle atmosphere will accompany tropospheric warming related to enhanced greenhouse gas emissions (Rind *et al.*, 1998; Roble and Dickinson, 1989). While there is already some observational evidence of cooling in the stratosphere and mesosphere of the northern hemisphere (Aikin *et al.*, 1991; Hauchecorne *et al.*, 1991; Keckhut *et al.*, 1999), it has been difficult to determine the relative importance of anthropogenic and natural forcing.

The CIRA-86 climatological reference model (Barnett and Chandra, 1990; Fleming *et al.*, 1990) is regarded as providing the best available global

representation of mean temperatures for the middle atmosphere. A convenient representation of this is model is publicly available in the MSISE-90 model (Hedin, 1991). However, the database from which the temperature model was synthesised is based largely on data for the northern hemisphere, and there is much scope for validation at southern hemisphere sites.

In this paper, temperature measurements of the stratopause region obtained with a Rayleigh lidar operated over a two year period at Kingston, Tasmania (43.0°S, 147.3°E) are compared with the MSISE-90 model.

3.2 Instrumentation

The Australian Antarctic Division and the University of Adelaide are developing an atmospheric lidar for investigating climate change in the middle atmosphere. The design of the instrument and its scientific program have been outlined by Morris *et al.* (1995) and Klekociuk *et al.* (1994). From 2001, the lidar will be located at Davis station in Antarctica (68.6°S, 78.0°E) where it will be used to profile temperature, wind velocity and aerosol loading in the stratosphere and mesosphere over a full solar cycle.

The first atmospheric measurements using the lidar were conducted at the University of Adelaide's field station at Buckland Park in 1992 and 1993. These measurements demonstrated the ability to undertake Rayleigh soundings using a novel optical arrangement which provided coaxially-aligned fields-of-view for the transmitter and receiver sections of the instrument (Argall and Jacka, 1996). A similar coaxial arrangement will be used at Davis to profile the wind field by measuring radial velocity components of the wind along various lines-of-sight.

During 1994, the lidar was relocated to the Australian Antarctic Division's headquarters at Kingston. There the instrument was upgraded initially to operate with a higher power laser and subsequently to incorporate a Fabry-Perot spectrometer in the receiver for Doppler wind and temperature measurements. The upgrade of the laser system required major alterations to the transmitter optics owing to the significantly higher energy density of the outgoing laser pulses. In order to undertake scientifically useful observations at Kingston while the new optics were being developed, an interim biaxial configuration of the instrument was devised which required the minimal amount of new hardware to be constructed. In this arrangement, the receiver hardware used at Buckland Park was retained, while new and independent optics was provided for the transmitter.

The optical system is shown schematically in Figure 1, and its main parameters are summarised in Table 1. Pulses of second harmonic (λ = 532 nm) light from a commercial Nd:YAG laser were transmitted vertically into the sky via a small beam expanding telescope. A detector placed behind beam steering mirror M1 was used to obtain a proxy

reading of the transmitted laser power. The Rayleigh backscatter from the sky was collected by a one metre aperture Cassegrain telescope and relayed through a narrow-band interference filter to a cooled photomultiplier. Two beam steering prisms in the optical relay system allowed the telescope to be moved in azimuth and zenith angle while maintaining optical alignment, although in practice movements of the telescope were small and only made at the start of each observing session to align the receiver to the beam of the transmitter. The pulses output by the photomultiplier were binned by a range-gated counter and accumulated synchronous with the firing of the laser. Backscattered light from ranges below about 15 km was prevented from reaching the photomultiplier by the action of two rotating shutters known as the Mirror Shutter and the Blanking Shutter. These shutters blocked the most intense levels of backscatter which would otherwise have resulted in significant pulse pileup and secondary pulse emission effects in the photon counting system.

Table 1. *Main parameters of the lidar system.*

Parameter	Value
Laser Wavelength	532 nm
Laser Pulse Energy	0.6 J
Laser Pulse Repetition Rate	50 Hz
Laser Pulse Duration (FWHM)	7 ns
Estimated Efficiency of Transmitting Optics	0.94
Estimated Efficiency of Receiving Optics	0.28
Interference Filter Bandwidth (FWHM)	0.3 nm
Estimated PMT Quantum Efficiency	0.14
Nominal Divergence of Transmitted Beam (full angle)	0.1 mrad
Nominal Field-of-View of Receiver (full angle)	0.5 mrad
Baseline Between Transmitter and Receiver	1 m
Telescope Collecting Area	0.79 m^2

An important requirement in the setup of the optical system was to ensure that the field-of-view of the receiver fully overlapped the field illuminated by the transmitter over the full range interval of interest, and that the two fields did not significantly move relative to each other over the course of each observing session. Significant misalignment will introduce a range-dependent efficiency term in the backscatter intensity as a function of range. Potentially, this effect can be very difficult to

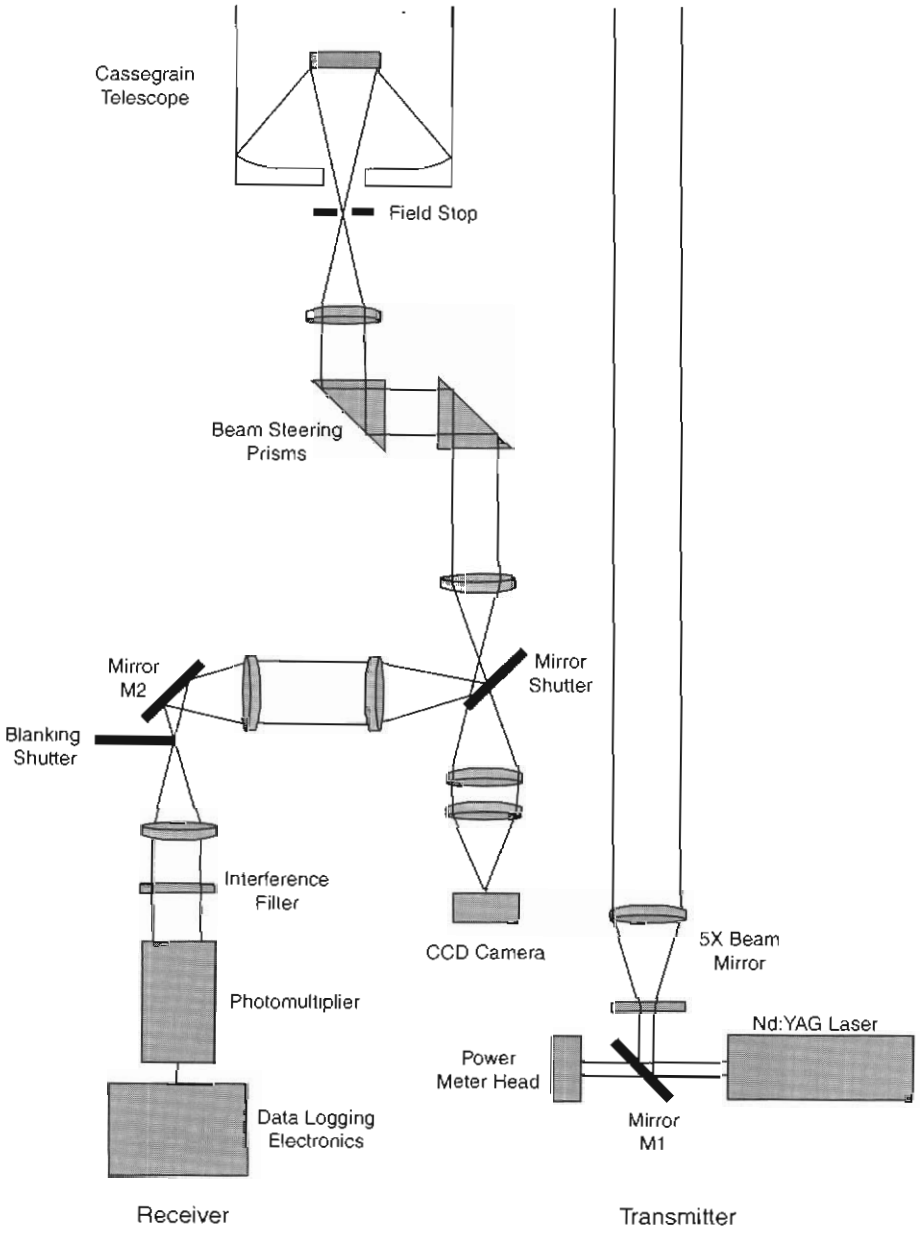


Figure 1. *Schematic of the biaxial optical system used for the observations.*

accurately account for, and is a significant issue for all lidar systems. In the case of the current system, ray tracing simulations indicated that adequate overlap would be expected beyond 10 km range given nominal tolerances for the performance of the optical elements and their alignment with respect to each other. The method used to verify the alignment of the system is outlined in the following section.

3.3 Data analysis

3.3.1 Observations

Scientifically useful observations with the lidar system were conducted on 22 nights between March 1997 and April 1999. Other observing sessions were conducted over this time, but the presence of cloud or the testing of equipment resulted in sparse data that were not suitable for detailed analysis. The range resolution was 333 m prior to 25 May 1997 and 94 m subsequently. All data were collected with an accumulation interval of 30 s. Prior to July 1997, the mechanical design of the Mirror Shutter and Blanking Shutter systems effectively limited the maximum detection altitude to 65 km. This was subsequently increased to about 110 km.

At the start of each observing session, the alignment of the two optical axes of the system was verified. This involved making small adjustments to the tilt of the telescope to maximise the backscatter return for a particular reference altitude, which was chosen to be approximately 44 km. Further optimisation of the alignment was made by through small azimuth and zenith angle tilts of mirror M1. A CCD camera was positioned in the return path below the Mirror Shutter, and provided images of the receiver field-of-view. By virtue of the design of the Mirror Shutter, it was possible to mechanically gate the backscattered light from a selected 15 km altitude interval through to the camera, and integrate the image for a suitable period. These images confirmed that the angular cross-section of the transmitted beam was smaller than the field-of-view of the receiver.

3.3.1 The lidar equation

For the instrumental configuration used in the present study, the number of photon counts N associated with altitude z can be expressed as

$$N(z) = C\,\rho(z)/z^2 + B \qquad (1)$$

where C is a proportionality constant that accounts for various instrumental and atmospheric parameters that remain constant during a measurement, ρ is the atmospheric density and B is a term describing the 'background' contribution from non-laser light and photomultiplier dark counts. Equation (1) is a simplified form of the 'lidar equation' for which it is assumed that the atmosphere is fully mixed, that aerosol scattering is negligible, and that the two-way atmospheric transmission is independent of altitude. Based on simulations of atmospheric transmission (Klekociuk, 1996), and analysis of scattering ratio profiles presented in Section 3.3, Equation (1) is deemed to hold for altitudes above 34 km in the present study.

An important assumption inherent in Equation (1) is that there are no significant altitude-dependent instrumental effects. This is deemed to be satisfied at least over the altitude range 25 km to 65 km based on the following considerations. Firstly, tests of the alignment outlined in Section 3.1 indicate that the fields-of-view of the receiver and the transmitter fully overlap for altitudes above 10 km. Secondly, the throughput of the optical system was found to vary by no more than 1% based on range-gated measurements of light from the twilight sky. Thirdly, the photon counting system was found to be linear to within 1% for the count rates typically encountered above 25 km (i.e. count rates less than 10 MHz) using an evaluation technique described by Donovan *et al.* (1993).

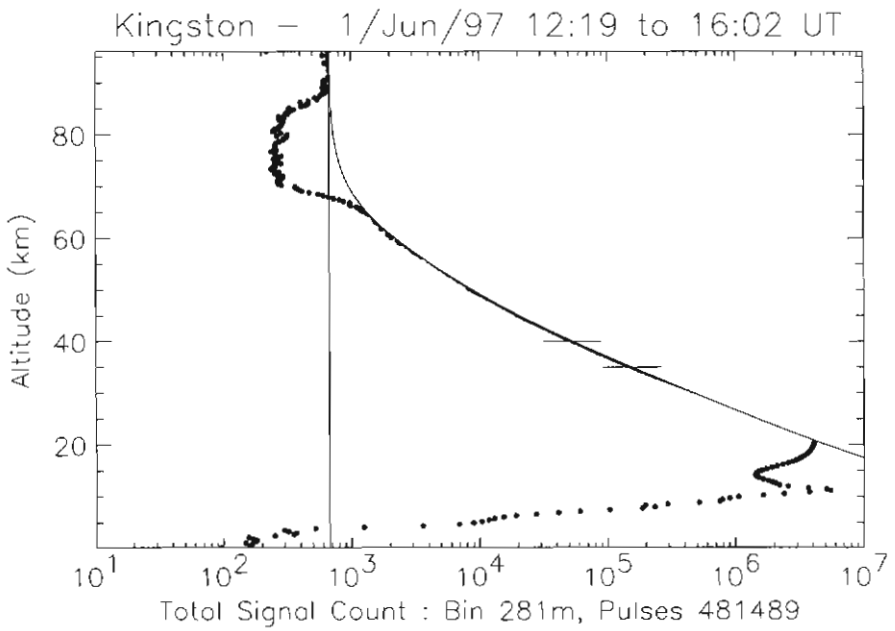


Figure 2. *Example of a typical lidar signal profile. The vertical bin width is 281 m, and the CSR interval is [35, 40] km. The data represent approximately 2.8 hours of 'live time' (see text), or approximately 504 000 laser shots. The observed data are represented by dots. Below approximately 10 km, the signal is attenuated by the combined effects of the Blanking Shutter and Mirror Shutter. Between about 10 km and 25 km, the data exhibit the effects of pulse pile-up. The dip between approximately 70 km and 85 km is due to the closing of the Mirror Shutter, and the data in this region are indicative of the photomultiplier dark count. The vertical line marks the estimated mean background count. The curved line is a fit of the lidar equation (Equation (1)) to the data over the vertical range of the 'clear signal region' (marked by the two horizontal ticks on the curve).*

3.3.2 *Density profile*

The following procedure was used to obtain profiles of proxy atmospheric density from the lidar data. Firstly, a merit factor was calculated from the product of the equivalent count rate for 45 km and the proxy laser power. Those profiles having an anomalously low value for the merit factor were discarded from further analysis. This generally related to data influenced by cloud. The remaining profiles were then co-added, and the background value (B in Equation (1)) was obtained by averaging the data from the upper altitude bins that were obviously unaffected by backscatter or the rotating shutters.

The altitude interval between 35 km and 40 km was defined as the 'clear signal region' or CSR. This region was selected to lie above the aerosol layer, and to contain data with a high signal-to-noise ratio. The atmospheric density $\rho_m(z)$ corresponding to the central altitude of each data bin within the CSR was then obtained from the MSISE-90 model.

The proportionality constant C in Equation (1) was solved by a weighted least-squares fit using the values of B, ρ and N over the clear signal region. The weighting function was given by

$$W(z) = 1/(N + [\varepsilon C\rho_m(z)/z^2]^2)$$

where ε is the estimated uncertainty in the model density. Using information in Labitzke *et al.* (1990), it was estimated that $\varepsilon = 0.02$ over the clear signal region. A typical signal profile and the resulting fit is presented in Figure 2. By rearranging Equation (1), proxy values of atmospheric density $\rho'(z)$ for the full altitude range were obtained. In the calculations, it was assumed that the mean molecular weight of the air was 28.964 (kg $kmol^{-1}$), and the gas constant was 8.3143 J K^{-1} mol^{-1}. For the analysis that follows, the proxy density data were binned into 281 metre increments (by combining consecutive 3 bin intervals). This gave an adequate compromise between measurement uncertainty and spatial resolution.

3.3.3 *Scattering ratio profile*

A useful diagnostic indicator of aerosol loading is the scattering ratio S, which may be defined as

$$S = \rho'(z)/\rho_m(z).$$

Values of S significantly greater than unity are indicative of enhanced scatter due to aerosols. An example profile is shown in Figure 3. Note that in this example, the data below about 25 km are influenced by pulse pile-up. Because of the way that the scattering ratio is defined, its mean value across the clear signal region is unity. All of the data collected during the study interval showed evidence of increasingly enhanced scatter below an altitude of about 34 km. There was some indication of structure in the form of 'ledges' at particular altitudes that persisted between some of the more closely spaced observations.

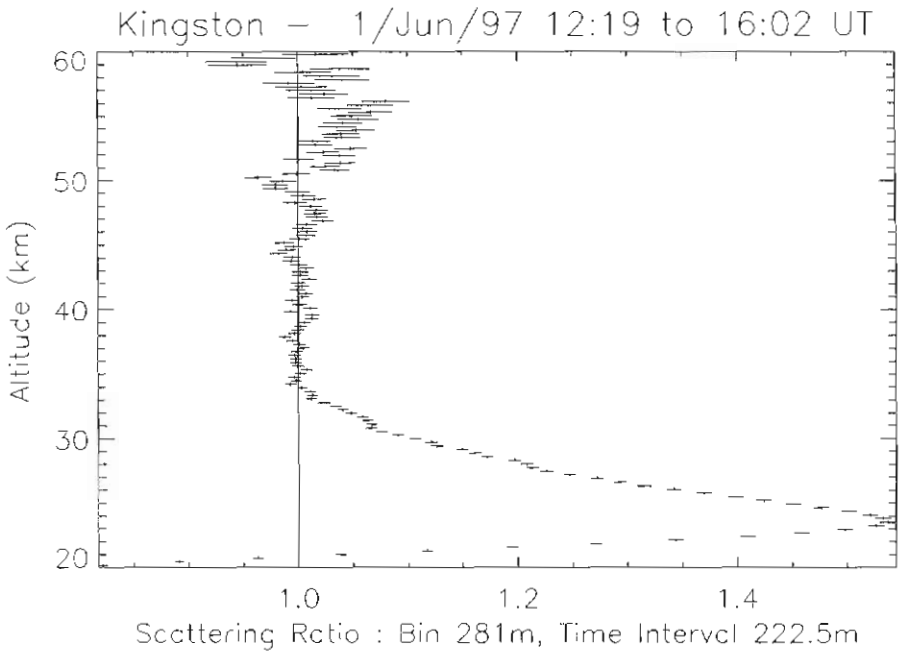


Figure 3. *Example of a typical scattering ratio profile. The data from Figure 2 have been used to calculate the ratio between the inferred density profile and the model. Pulse pile-up below about 25 km results in the under estimation of the scattering ratio. The data between 25 km and 35 km are indicative of enhanced scattering due to the presence of aerosols.*

3.3.4 Temperature profile

Derivation of temperatures above the aerosol layer proceeded on the assumption that the atmosphere is in hydrostatic equilibrium (requiring that turbulence is unimportant for the spatial resolution and averaging time of the measurements) and that the atmospheric gases obeyed the perfect gas law. The standard lidar temperature retrieval algorithm is iterative, and works down in altitude from a reference level (the 'seeding' altitude) where the pressure or temperature are known or assumed. For each altitude bin, the weight of the overlying atmosphere is calculated, and then combined with the associated density estimate to yield temperature by invoking the perfect gas law. A detailed description of the method can be found in Hauchercorne and Chanin (1980). For the calculations presented here, it was assumed that the gravitational acceleration was 9.807 m s^{-2} at the observatory, and follows an inverse-square law with distance from the centre of the Earth.

In the present case, the retrieval algorithm was initialised using temperatures from the MSISE-90 model. The choice of seeding altitude has a bearing on the convergence of the method. Convergence will be slow if the statistical uncertainty of the proxy density measurements near the seeding altitude is high. Through experimentation involving realistic variation of both the seeding altitude and the reference temperature, it was found that stable temperature profiles could be obtained below an altitude of 55 km using a seeding altitude of 60 km in all cases. It was also found that temperatures over the altitude range 35 km to 60 km were not significantly influenced by the choice of CSR within this range.

An example temperature profile is shown in Figure 4. Note that the retrieved temperature profile appears to track the model predictions reasonably well. Apparent also are variations from the model of up to about 5 K in amplitude. These are discussed in Section 3.4.2. Below about 35 km, the observations diverge from the model due to aerosol contamination. Unfortunately, it was not possible to directly compare radiosonde temperatures with any of the lidar temperature estimates owing to the fact that all of the balloon flights terminated below 30 km.

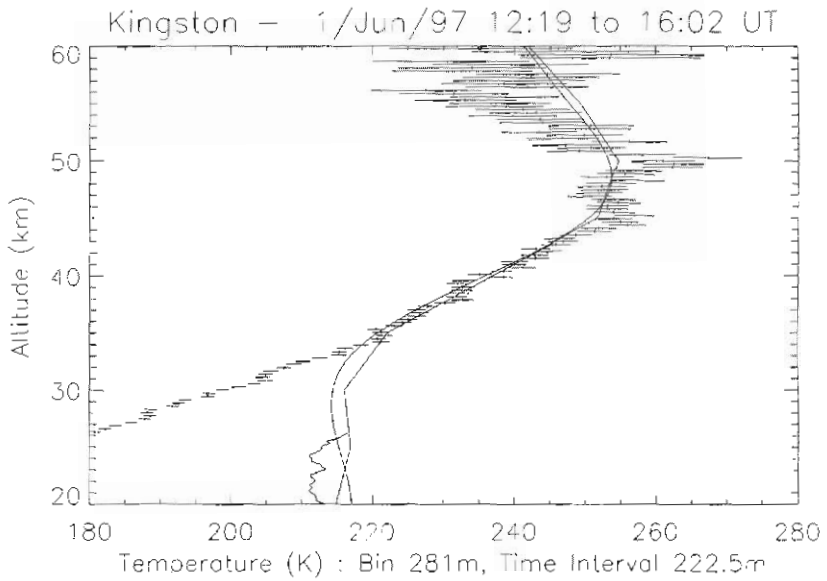


Figure 4. *Temperature profile retrieved from the data of Figure 2. The retrieval algorithm was initiated with a model temperature value for 60 km. Error bars are plotted as plus and minus 2 standard deviations. The jagged continuous line below about 26 km is radiosonde temperature data from the 1100 UT balloon flight at Hobart Airport. The continuous curve overlaying the data is the MSISE-90 model, while the piecewise curve is the CIRA-86 model.*

3.4 Discussion

3.4.1 Mean stratopause temperatures

The MSISE-90 model predicts a seasonal amplitude of about 20 K at the stratopause above Kingston. The model also predicts that the height of the 'knee' in the temperature profile should not change by more than about 1 km during the year.

In order to examine the annual temperature at the stratopause in the observational data, daily temperature profiles were produced with 1 km altitude binning, and weighted temperature averages were calculated over the interval 45 km to 50 km (which straddles the mean height of the stratopause in the MSISE-90 model). The resulting daily values are plotted as a function of time of year in Figure 5 and are summarised in Table 2.

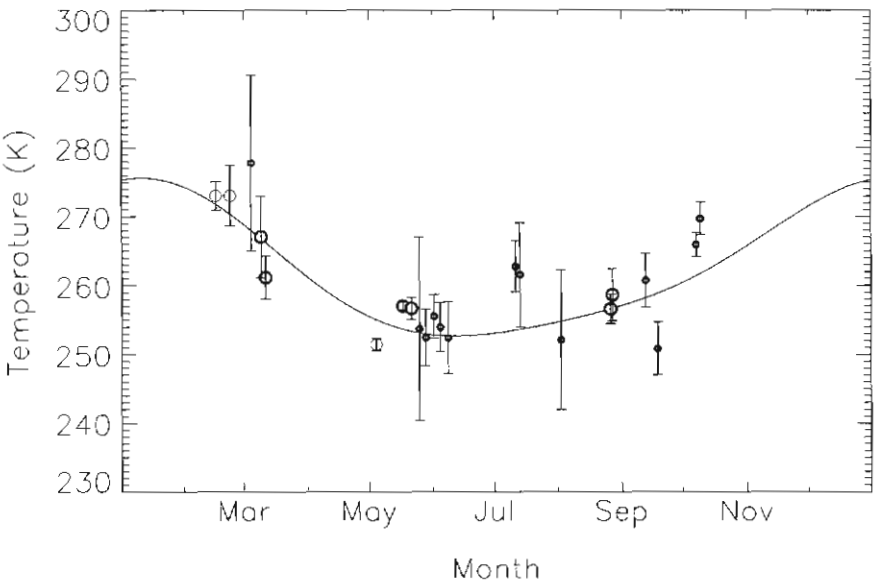


***Figure** 5. Mean stratopause region temperatures. The data from each observing session were binned to a single profile having an altitude interval of approximately 1 km. The temperature profile was then obtained using a CSR interval of [35, 40] km and then averaged over the interval [45, 50] km. Data for 1997 are plotted with small bold circles, data for 1998 have large bold circles, and data for 1999 have open circles. The error bars have a range of plus and minus one standard deviation in the mean. The solid curve represents the temperature from the MSISE-90 model for each day of the year at local midnight averaged over the interval [45, 50] km.*

Table 2. *Mean stratopause region temperatures from the observations and the MSISE-90 model. The observed temperatures were derived using a bin width of approximately 1 km and a CSR interval of [35, 40] km, and then averaged over the interval [45, 50] km. The 'live time' refers to the equivalent amount of continuous observing time used to produce each observed temperature estimate. To convert 'live time' to the number of laser shots, multiply by 180 000. The model temperatures are averages over the altitude interval [45, 50] km for the central time of each observing session. The error associated with each value is one standard deviation in the mean across the averaging region.*

Central Time of Observing Session (UT)	Observed Temperature (K)	MSISE-90 Model Temperature (K)	Live Time (hours)
1997 May 04 1535	277.8 ±12.8	268.2 ±1.1	7.8
1997 May 25 1109	253.8 ±13.3	253.9 ±0.4	0.2
1997 May 28 1521	252.5 ±4.1	253.6 ±0.4	4.2
1997 Jun 01 1410	255.6 ±3.2	253.4 ±0.4	2.8
1997 Jun 04 1421	254.0 ±3.6	253.4 ±0.5	2.3
1997 Jun 08 1200	252.4 ±5.2	253.4 ±0.6	2.5
1997 Jul 11 0855	262.7 ±3.7	254.3 ±0.6	1.0
1997 Jul 13 1253	261.6 ±7.6	254.1 ±0.5	1.6
1997 Aug 02 1038	252.1 ±10.1	255.3 ±0.3	1.4
1997 Sep 12 1638	260.7 ±3.9	258.3 ±0.9	2.9
1997 Sep 18 1404	250.9 ±3.8	258.9 ±1.0	2.9
1997 Oct 07 1245	265.9 ±1.7	261.4 ±1.1	4.6
1997 Oct 09 1246	269.7 ±2.4	261.7 ±1.1	4.9
1998 Mar 09 1201	267.1 ±5.9	267.4 ±0.9	0.9
1998 Mar 11 1539	261.2 ±3.1	266.7 ±1.0	5.3
1998 May 17 1241	257.0 ±0.6	254.3 ±0.3	3.3
1998 May 21 1308	256.7 ±1.6	254.0 ±0.3	2.1
1998 Aug 26 1538	256.6 ±2.1	256.8 ±0.6	9.2
1998 Aug 27 1246	258.6 ±3.8	256.9 ±0.6	5.5
1999 Feb 15 1244	273.0 ±2.1	272.3 ±0.7	1.9
1999 Feb 22 1027	273.0 ±4.4	270.9 ±0.8	1.2
1999 May 04 1314	251.5 ±0.9	255.6 ±0.3	4.1

About half the range in the statistical uncertainty of these data can be ascribed to systematic variations across the averaging region. From examination of the profiles it was noted that the temperature inversion at the stratopause tends to be sharper in autumn and spring which is consistent with the prediction of the model. It is apparent that the data and model track reasonably well given the measurement precision. Spearman's rank correlation test on the observed and model data gave a correlation coefficient of 0.51 for 21 degrees of freedom, which indicates the probability that the two data sets are related purely by chance is 7%. Of additional note is the consistency between temporally adjacent data points. An example of this can be seen over the interval 28 May to 4 June 1997.

3.4.2 Temperature perturbations

It was observed that the daily temperature profiles over the altitude range 35 km to 55 km exhibited statistically significant perturbations with amplitudes of up to 5 K and apparent vertical scales of a few kilometres. This behaviour can be seen in Figure 4. At higher altitudes, similar variations would tend to be masked by the level of measurement uncertainty and the influence of convergence in the temperature retrieval method. The observed features did not appear to be stationary in altitude between observing sessions, although in most cases several days spanned adjacent sessions. These variations could be related to the propagation of gravity waves, planetary waves and tides, although the detailed analysis of these features is beyond the scope of this paper.

3.5 Conclusions

The analysis presented here is at a preliminary stage, but provides some indication as to the state of the middle atmosphere above Kingston. Firstly, aerosol scattering is evident up to an altitude of about 34 km. Secondly, observed mean temperatures at the stratopause are consistent with predictions of the MSISE-90 model over the autumn to spring months. Thirdly, temperature perturbations of up to about 5 K in amplitude were apparent between the top of the aerosol layer and 55 km altitude. The perturbations may be related to the propagation of atmospheric waves, although further quantitative analysis is required.

At Davis, intensive Rayleigh lidar observations will be conducted using both coaxial and biaxial optical configurations. It is planned to fly radiosondes to altitudes of up to 35 km, which when combined with the expected lower ceiling of the stratospheric aerosol layer, should allow the density retrieval method to be calibrated with *in situ* measurements. In addition, temperature profiles will also be obtained from incoherent Doppler and Raman scatter measurements, which will provide for important cross-comparisons potentially extending down through the aerosol layer.

3.6 Acknowledgements

I am grateful to J.L. Innis for assistance in operating the lidar during collection of the observations. I thank members of the Atmospheric and Space Physics group, the Science Technical Support section and Engineering Branch of the Australian Antarctic Division for their involvement in the development of the lidar, in particular R.J. Morris, D.J. Rasch, D.J. Murphy, A. Fleming and G. Ratcliffe. D.J. Watts provided assistance with software development. B. Copplestone of the Australian Bureau of Meteorology supplied the radiosonde data from Hobart Airport. P.S. Argall and D. Creighton undertook earlier work in the development of the lidar at the University of Adelaide. The Australian Antarctic Division, the University of Adelaide, and the Australian Research Grants Scheme have provided funding for the lidar project.

References

Aikin, A.C., Chanin, M.L., Nash, J. and Kendig, D.J. (1991). Temperature trends in the lower mesosphere. *Geophysical Research Letters*, 18 (3): 416–419.

Argall, P.S. and Jacka, F. (1996). High-pulse-repetition-frequency lidar system using a single telescope for transmission and reception. *Applied Optics*, 35 (15): 2619–2629.

Barnett, J.J. and Chandra, S. (1990). COSPAR international reference atmosphere. Chapter 1, Atmosphere grand mean, *Advances in Space Research*, 10 (12): 7–10.

Donovan, D.P., Whiteway, J.A. and Carswell, A.I. (1993). Correction for nonlinear photon-counting effects in lidar systems. *Applied Optics*, 32 (33): 6742–6753.

Fleming, E.L., Chandra, S., Barnett, J.J. and Corney, M. (1990). COSPAR international reference atmosphere. Chapter 2, Zonal mean temperature, pressure, zonal wind, and geopotential height as a function of latitude. *Advances in Space Research*, 10 (12): 11–59.

Hauchecorne, A. and Chanin, M.-L. (1980). Density and temperature profiles obtained by lidar between 35 and 70 km. *Geophysical Research Letters*, 7 (8): 565–568.

Hauchecorne, A., Chanin, M.-L. and P. Keckhut, P. (1991). Climatology and trends of the middle atmospheric temperature (33–87 km) as seen by Rayleigh lidar over the south of France. *Journal of Geophysical Research*, 96 (D8): 15297–15309.

Hedin, A.E. (1991). Extension of the MSIS thermosphere model into the middle and lower atmosphere. *Journal of Geophysical Research*, 96 (A2): 1159–1172.

Keckhut, P., Schmidlin, F.J., Hauchecorne, A. and Chanin, M.L. (1999). Stratospheric and mesospheric cooling trends estimated from U.S. rocketsondes at low latitude stations (8°S–34°N), taking into account instrumental changes and natural variability. *Journal of Atmospheric and Solar-Terrestrial Physics*, 61: 447–459.

Klekociuk, A.R. (1996). Extension of density profiles obtained by Rayleigh lidar into the lower stratosphere using Raman scattering. In: Morris, R.J. (Ed.). *ANARE Research Notes Number 95*. Australian Antarctic Division. Pp. 296–310.

Klekociuk, A.R., Argall, P.S., Morris, R.J., Yates, P., Fleming, A., Vincent, R.A., Reid, I.M, Greet, P.A. and Murphy, D.J. (1994). The Australian Antarctic lidar facility. In: Wang, J. and Hays, P.B. (Eds.). *Proceedings SPIE 2266 Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research*. The International Society for Optical Engineering. Pp. 624–634.

Labitzke, K., Naujokat, B. and Bernet, J.J. (1990). Interannual variability. *Advances in Space Research*, 10 (12): 163–168.

Morris, R.J., Monselesan, D.P. and Klekociuk, A.R. (1995). Australian Antarctic middle and upper atmosphere physics – a new direction. *Advances in Space Research*, 16: 151–162.

Rind, D., Shindell, D., Lonergan, P. and Balachandran, N.K. (1998). Climate change and the middle atmosphere. Part III: The doubled CO_2 climate revisited. *Journal of Climate*, 11: 876–894.

Roble, R.G. and Dickinson, R.E. (1989). How will changes in carbon dioxide and methane modify the mean structure of the mesosphere and thermosphere? *Geophysical Research Letters*, 16 (12): 1441–1444.

4. DERIVATION OF GRAVITY-WAVE SPECTRA FROM RAYLEIGH LIDAR OBSERVATIONS: A REVIEW OF THE FORMALISM AND A COMPARISON WITH SOME PUBLISHED ANALYSES

J.L. Innis, A.R. Klekociuk and D.J. Murphy

Atmospheric and Space Physics
Australian Antarctic Division
Kingston Tasmania 7050
(email: john.innis@aad.gov.au)

Abstract

This paper presents the mathematical background of the derivation of atmospheric gravity-wave spectra from Rayleigh lidar observations of fractional density and/or temperature fluctuations. This is accomplished by first considering the relationships between the observed wave-induced density or temperature change and the quantities that can be derived from this (horizontal wind variance, energy per unit mass measures) for monochromatic waves, using the 'polarisation relations' given by Hines (1960). The relationships between the observed and derived quantities for single waves are then generalised to the calculation of gravity wave spectra.

Specifically, the relationship between the wave-induced perturbation in density and horizontal wind is derived. Attention is given to the validity of two approximations used to derive this relationship – that of limiting to 'small vertical scale' waves (i.e. large vertical wave-number waves only) and the 'hydrostatic approximation' (allowable wave frequencies only up to ~ 0.3 times the Brunt-Vaisala frequency), and also to the conditions necessary for the wave-induced relative density fluctuations to be equal in magnitude to the relative temperature fluctuations. These are shown to be valid for the likely range of waves observable with a Rayleigh lidar covering the stratosphere and mesosphere. The relationship between the temperature (or density) fluctuations and wave energy per unit mass is also presented.

The derivation of gravity wave spectra is then considered. An often used technique in the analysis of lidar data is to derive vertical wave number spectra of wave-induced horizontal wind variance calculated from the wave-induced fractional density fluctuations. However, for 'small-scale waves' where the only density change is due to vertical motion of air parcels, the density changes are greatest for the higher frequency waves which have the least horizontal velocity motion. A better interpretation of the scaled squared density fluctuations is in terms of the 'potential

energy per unit mass' spectrum, which is valid for all wave frequencies. A comparison with published gravity wave analysis is made, and some apparent inconsistencies between various methods are addressed. While much of the material presented here is available in a number of sources (cited below), this paper attempts, in tutorial style, to collect the relevant material into a relatively complete form for future reference by the lidar group at the Australian Antarctic Division.

4.1 Introduction

Rayleigh lidars have been used for a number of years to observe the temperature and dynamics of the lower and middle atmosphere. The signal backscattered from altitudes above about 35 km (above aerosol contamination), when corrected for an inverse-square range dependence, is effectively proportional to the local density. Assumption of hydrostatic equilibrium and the ideal gas law enables temperatures to be calculated (e.g. Hauchecorne and Chanin, 1980). Short-term fluctuations in density and temperature have been ascribed to the presence of atmospheric gravity waves. Mathematical techniques have been developed to use the measured fractional fluctuations in density and temperature (i.e. the change in density or temperature at a given height divided by the mean density or temperature at this height) to produce estimates of the gravity wave spectra in the middle atmosphere (e.g. Mitchell *et al.*, 1991; Senft *et al.*, 1993; Whiteway and Carswell, 1994; Meriwether *et al.*, 1994, among others).

The Atmospheric and Space Physics section of the Australian Antarctic Division (AAD) is currently assembling a high power Rayleigh/Doppler lidar for deployment at Davis station, Antarctica, in 2000. A description of this instrument is given in Klekocuik *et al.* (1994). At the time of writing test observations are underway at AAD headquarters in Kingston, Tasmania. These data are being used to evaluate instrument performance, to assist in the development of data reduction and analysis programs, and also to obtain scientifically valuable measurements of the atmosphere at a southern middle-latitude site (Klekociuk, 2001).

As part of the program, a review of techniques used in the analysis of lidar data was undertaken. This paper summarises aspects that are relevant for the derivation of gravity wave spectra from lidar data. In what follows a basic familiarity with the lidar technique and gravity waves will be assumed.

4.2 Definition of symbols used

A list of the main symbols used is given below.

U' = (total) horizontal perturbation velocity due to a gravity wave. In most of the following the perturbation will be assumed to be entirely in the x-direction, through judicious choice of axes.

W' = vertical perturbation velocity due to a gravity wave.

λ_x = horizontal wavelength of the gravity wave in the x-direction.

λ_z = vertical wavelength of the gravity wave.

k = horizontal wave number in x-direction = $2\pi/\lambda_x$. (Sometimes denoted as k_x in the literature).

l = horizontal wave number in y-direction = $2\pi/\lambda_y$. (Sometimes denoted as k_y in the literature).

Note: In most of the following the wave will be assumed to propagate only in the x-direction, so that the total horizontal wave number $k_H = k$, i.e. $l = 0$, and the total horizontal wavelength $\lambda_T = \lambda_x$. Where this is not the case then $k_H^2 = k^2 + l^2$, and $\lambda_T^2 = \lambda_x^2 + \lambda_y^2$.

m = vertical wave number = $2\pi/\lambda_z$. Sometimes denoted as 'k_z' in the literature. m (italic m) is used to avoid confusion with 'm', the symbol for metre.

ω = wave frequency.

v = horizontal phase speed of the wave = ω/k.

ρ = atmospheric density, normally implicitly $\rho(z, t)$, atmospheric density at a given height and time.

ρ_0 = mean density at a given height.

ρ' = normalised (fractional) density perturbation = $(\rho(z, t) - \rho_0)/\rho_0$ = $\Delta\rho/\rho_0$.

T, T_0 = atmospheric temperature, and mean temperature, as for density.

T′ = normalised temperature perturbation, = $\Delta T/T_0$, etc., as for density.

p, p' = atmospheric pressure and normalised pressure perturbation respectively, $p' = \Delta p/p_0$ as for density etc.

γ = ratio of specific heats, approximately 1.4 in the middle atmosphere.

g = acceleration due to gravity, about 9.5 m s^{-2} in the middle atmosphere.

C = speed of sound, about 300 m s^{-1} in the middle atmosphere.

i = $\sqrt{(-1)}$.

n = number density, the number of atoms or molecules per unit volume.

k = Boltzmann's constant.

f = inertial frequency, $f = 2\Omega \sin\varphi$, where Ω is the angular rotation frequency of the Earth, and φ is the site latitude.

ω_a = acoustic cut off frequency, the lowest frequency acoustic oscillation the atmosphere can support, $\omega_a = \gamma g/2C$. About 2.2×10^{-2} rad s^{-1} in the middle atmosphere.

N = Brunt-Vaisala frequency, the highest frequency buoyancy oscillation the atmosphere can support, $N = (\gamma - 1)^{0.5} g/C$. About 2.0×10^{-2} rad s^{-1} in the middle atmosphere. N is slightly less than ω_a over most of the middle atmosphere. Note: the values of γ, g, and C set the value of N.

$F_\rho(m)$ = vertical wave number spectrum of the normalised density perturbation, ρ' (sometimes $E_\rho(m)$ in the literature, and similarly for the following as well).

$F_T(m)$ = vertical wave number spectrum of the normalised temperature perturbation, T'.

$F_U(m)$ = vertical wave number spectrum of the horizontal velocity perturbation, U'.

$F_W(m)$ = vertical wavenumber spectrum of the vertical velocity perturbation, W'.

$F_\rho(\omega)$ = angular frequency spectrum of the normalised density perturbation, ρ'.

$F_T(\omega)$ = angular frequency spectrum of the normalised temperature perturbation, T'.

$F_U(\omega)$ = angular frequency spectrum of the horizontal wind, U'.

4.3 Formulation of the problem

The presence of gravity waves in the lower and middle atmosphere is normally demonstrated from observations of spectra of some perturbation quantity, such as horizontal wind or atmospheric temperature. Gravity wave spectra are expected to show a dependence on both the scale (e.g. horizontal and vertical wavelengths) and the frequency of the individual waves present at the time of observation. (An azimuthal dependence is also to be expected, but this will be neglected as azimuthal symmetry will be assumed unless otherwise stated.) Horizontal wavenumber, vertical wavenumber, and wave frequency are related by the gravity wave dispersion relation (see below), so that specifying any two of these quantities determines the third for a given wave. The basic mathematical formalism of atmospheric gravity waves can be found in Hines (1960).

The gravity wave dispersion equation is (Hines, 1960):

$$(\omega^2 - \omega_a^2)\omega^2/C^2 - \omega^2(m^2 + k^2) + N^2k^2 = 0. \quad (1)$$

This can be written as $m^2 = k^2(N^2 - \omega^2)/\omega^2 + (\omega^2 - \omega_a^2)/C^2$, and as $k = \omega/v$ (from the symbol definitions above), we can write:

$$m^2 = (N^2 - \omega^2)/v^2 - (\omega_a^2 - \omega^2)/C^2. \quad (2)$$

Alternatively, solving for k^2, we have,

$$k^2 = \omega^2/(N^2 - \omega^2)(m^2 + (\omega_a{}^2 - \omega^2)/C^2). \quad (3a)$$

We will later make the assumption that $m^2 >> (\omega_a{}^2 - \omega^2)/C^2$, so that the dispersion relation becomes:

$$k^2(N^2 - \omega^2) = m^2\omega^2. \quad (3b)$$

These will be used in the following sections.

Note that these equations do not take into account the effect of the rotation of the Earth. The dispersion relation as given in Equation (1) allows for gravity waves of frequencies from $\omega = 0$ up to $\omega = N$. In fact, in the real atmosphere, the lowest (intrinsic) frequency a gravity wave can have is $\omega = f$, where f is the inertial frequency ($f = 2\Omega \sin\varphi$) as defined in Section 4.2. For gravity waves of very low frequency (say $\omega \leq 3f$) the effect of the rotation of the Earth is to modify the dispersion and polarisation equations, as will be discussed further in a later section. Here we will just give the modified dispersion relation for comparison: $k^2(N^2 - \omega^2) = (\omega^2 - f^2)m^2$.

In the calculation of gravity wave spectra from ground based observations, normally the vertical wave number (m) spectra and frequency (ω) spectra are determined. The complication between intrinsic frequency and that observed from ground based stations is noted, but will not be discussed here. Also mentioned in passing is the effect that background winds may have on the vertical wave number spectra. The comprehensive paper by Eckermann (1995) is recommended for a treatment of this topic. It is usually assumed (e.g. VanZandt, 1982; Fritts and VanZandt, 1993) that although gravity wave properties depend on both m and ω, the spectra are separable in these quantities, i.e.

$$F_{GW}(m, \omega) = F_{GW}(m)F_{GW}(\omega),$$

where F_{GW} is the spectrum of a given property of the gravity wave - i.e. fluctuations in temperature, density, energy, etc. Under the assumption of spectrum separability, the vertical wave number and frequency spectra can be determined individually. In general, the data generally seem to indicate good evidence for a 'universal' spectrum of gravity waves, largely invariant in time and geographic location (e.g. VanZandt, 1982), with

$$F_{GW}(m) \propto m^{-q},$$

and

$$F_{GW}(\omega) \propto \omega^{-p},$$

with q around 2 or 3 for wavenumbers greater than a 'characteristic value', and around zero or -1 otherwise, and p is around 5/3. Note that $\int F_{GW}(\alpha)d\alpha = 1$, where α is m or ω, etc., and the integration is performed over all allowable frequencies, from normalisation to the total energy. (See e.g. Fritts, 1995, and references therein.)

Some earlier work on gravity wave spectra was performed using MF radars to determine the spectrum of horizontal wind perturbations due to the waves, and how these winds change with height. That is, the vertical wave number spectra of the horizontal velocity variance ($F_U(m)$) were determined. Rayleigh lidar observations produce density and temperature data, normally in a zenith-only direction. To make the lidar data comparable with that from radar the wave induced perturbations observed in density and temperature are sometimes transformed to a quantity comparable to the perturbation induced in the horizontal velocity. This can be done for monochromatic waves using the wave 'polarisation relations' derived by Hines (1960), which describe how the perturbations in pressure, density, horizontal velocity and vertical velocity relate to each other, viz.:

$$p'/P = \rho'/R = U'/X = W'/Z, \tag{4}$$

where

$$P = \gamma\omega^2[m - i(1 - \gamma/2)g/C^2], \tag{5a}$$

$$R = \omega^2 m + i[(\gamma - 1)gk^2 - \gamma g\omega^2/2C^2], \tag{5b}$$

$$X = \omega k C^2[m - i(1 - \gamma/2)g/C^2], \tag{5c}$$

$$Z = \omega[\omega^2 - k^2C^2]. \tag{5d}$$

Hence it is seen that the ratio of perturbations in the normalised density and the horizontal wind, for a given wave, can be relatively simply related as X/R, which gives:

$$U' = \rho'(\omega k C^2[m - i(1 - \gamma/2)g/C^2])/(\omega^2 m + i[(\gamma - 1)gk^2 - \gamma g\omega^2/2C^2]). \tag{6}$$

In most cases however individual waves are not observed, so that measurements of the quantities ω, k, and m are not available. Instead a spectrum of waves results. In this case, in order to simplify the expression for X/R, it is assumed that the wave frequency ω, is no greater than about 0.3 times the Brunt-Vaisala frequency. The resulting simplified equation is derived in Section 4.4, while the validity of the assumption is considered in Section 4.5.

It is also often assumed that the fractional changes in temperature, T', are equal in magnitude to the fractional changes in density, ρ'. The conditions required for this to be the case are considered in Section 4.6.

Also of interest is the 'energy per unit mass' of a gravity wave. For example kinetic energy per unit mass, KE_M, is clearly related to the sum of the squares of the horizontal and vertical velocity perturbations. (i.e. $KE_M = {}^1/_2\,(|U'|^2 + |W'|^2)$. For the case where W' is small compared with U' (low frequency waves) KE_M can be directly equated with the square of the horizontal winds. A satisfactory calculation of KE_M can be made from observations of the fractional temperature fluctuations, under most conditions. A calculation of potential energy per unit mass can also be

made in a similar way, and is valid for all wave frequencies. These aspects are discussed in Section 4.7 below.

Section 4.8 generalises the results for monochromatic waves to the derivation of gravity wave spectra. In these cases it is necessary to perform integrations over a range of ω in order to produce the required spectra. The relationships between the spectra of certain quantities (such as ρ' and U', or ρ' and KE_M for example) depend on which frequency ranges are being considered, as described in that section. It is demonstrated that it is important to choose a frequency bandwidth appropriate to the experimental parameters when estimating a derived spectrum (such as kinetic energy) from observed quantities (such as density fluctuations), and also to ensure that an appropriate choice of dispersion and polarisation equations is made – i.e. if very low frequency waves ($\omega \sim f$) are contributing to the spectra, then the effects of the Earth's rotation must be included in these equations. We consider that transforming the spectra of the wave-induced density (or temperature) fluctuations into potential energy per unit mass spectra, via a scaling proportional to $(g/N)^2$, is the preferred method of interpreting lidar data.

Section 4.9 makes a brief survey of how these mathematical techniques have been applied to the derivation of gravity wave spectra, as presented by several groups in the field. A number of apparent inconsistencies are addressed. It appears that in some instances inappropriate expressions have been used in the analysis, as explained in that section.

The appendices contain some mathematical detail omitted from the main body of the text.

4.4 Relation between density and horizontal wind fluctuations for monochromatic waves

As mentioned, a number of lidar groups present estimates of the spectrum of the wave induced horizontal wind fluctuations as derived from scaling observations of the wave induced density (or temperature) fluctuations. This section derives the relationship between horizontal winds and density changes for monochromatic gravity waves.

Proceeding from Equation (6) above, we calculate the squared magnitude of (X/R), which removes the complex terms, hence

$$|U'|^2 = |\rho'|^2|X|^2/|R|^2.$$

Working with the numerator first, we have:

$$|X|^2 = \omega^2 k^2 C^4 (m^2 + (1 - \gamma/2)^2 g^2/C^4).$$

Using Equation (1), the dispersion relation, we obtain:

$$|X|^2 = \omega^2[\omega^2/(N^2 - \omega^2)]\,(m^2 + (\omega_a^2 - \omega^2)/C^2)(C^4 m^2 + (1 - \gamma/2)^2 g^2). \quad (7a)$$

With the assumption that $\omega^2 << N^2$, and hence $\omega^2 << \omega_a^2$, this becomes

$$|X|^2 = (\omega^4/N^2)(m^2 + \omega_a^2/C^2)(C^4m^2 + (1 - \gamma/2)^2g^2). \quad (7b)$$

This is the approximation we obtain when neglecting the higher frequency waves. For $\omega = 0.3N$, using Equation (7b) instead of Equation (7a) gives an error of order 10%.

For $|R|^2$ we have:

$$|R|^2 = \omega^4m^2 + (\gamma - 1)^2g^2k^4 - 2\gamma(\gamma - 1)g^2k^2\omega^2/2C^2 + \gamma^2g^2\omega^4/4C^4.$$

Using $N^2 = (\gamma - 1)g^2/C^2$, and $\omega_a^2 = (\gamma\omega/2C)^2$, we have

$$|R|^2 = \omega^4m^2 + (\gamma - 1)^2g^2k^4 - \gamma N^2k^2\omega^2 + \omega^4\omega_a^2/C^2.$$

Substituting for k^2 using Equation (3a), we have,

$$|R|^2 = \omega^4m^2 + (\omega^4(\gamma - 1)^2g^2/(N^2 - \omega^2)^2)[m^2 + (\omega_a^2 - \omega^2)/C^2]^2 -$$
$$\gamma N^2(\omega^2/(N^2 - \omega^2))[m^2 + (\omega_a^2 - \omega^2)/C^2]\omega^2 + \omega^4\omega_a^2/C^2. \quad (8a)$$

Again, for $\omega^2 << N^2$, ω_a^2, noting that $(\gamma - 1)g^2 = (NC)^2$, and collecting like terms this becomes:

$$|R|^2 = \omega^4\{[m^2 + \omega_a^2/C^2] + (\gamma - 1)\ C^2/N^2[m^2 + \omega_a^2/C^2]^2 - \gamma[m^2 + \omega_a^2/C^2]\},$$

or

$$|R|^2 = \omega^4[m^2 + \omega_a^2/C^2]\ (\gamma - 1)\{C^2/N^2[m^2 + \omega_a^2/C^2]\ -1\}. \quad (8b)$$

Note that in Equations (7b) and (8b) the term $\omega^4[m^2 + \omega_a^2/C^2]$ appears.

Forming the ratio required, using Equations (7b) and (8b), gives:

$$|X|^2/|R|^2 = 1/N^2(C^4m^2 + (1 - \gamma/2)^2g^2)/[(\gamma - 1)\{C^2/N^2\ [m^2 + \omega_a^2/C^2]\ -1\}] \quad (9)$$

This expression can be simplified significantly. Taking out a factor of $(CN)^2$ in the numerator, and multipliying out in the denominator, gives:

$$|X|^2/|R|^2 = C^2(C^2/N^2m^2 + (1 - \gamma/2)^2g^2/C^2N^2)/[(\gamma - 1)(C^2/N^2m^2 + (\omega_a/N)^2 - 1)].$$

As $(g/CN)^2 = 1/(\gamma - 1)$, and $(\omega_a/N)^2 = \gamma^2/[4(\gamma - 1)]$, we have:

$$|X|^2/|R|^2 = C^2(C^2/N^2m^2 + (1 - \gamma/2)^2/(\gamma - 1))/[(\gamma - 1)(C^2/N^2m^2 + (\gamma^2 - 4\gamma + 4)/\{4(\gamma - 1)\}],$$

which yields directly:

$$|X|^2/|R|^2 = (C^2/(\gamma - 1))[C^2/N^2m^2 + (2 - \gamma)^2/\{4(\gamma - 1)\}]/[C^2/N^2m^2 + (\gamma - 2)^2/\{4(\gamma - 1)\}],$$

and hence:

$$|X|^2/|R|^2 = (C^2/(\gamma - 1)) = (g/N)^2. \quad (10)$$

Hence, subject to the condition of $\omega \leq 0.3N$, we have the simple expression:

$$|U'|^2 = (g/N)^2|\rho'|^2. \quad (11)$$

That is, the horizontal velocity variance can be related to the normalised density variance by simply scaling by $(g/N)^2$, for the conditions noted. This is the form used in a number of papers in the literature (e.g. Senft and Gardner, 1991; Senft *et al.*, 1993) to scale lidar density fluctuation observations to produce estimates of the vertical wavenumber spectra of the horizontal winds.

N and g are only approximately constant in the middle atmosphere, as they vary with height, as do U′ and ρ′. Thus Equation (11) should read

$$|U'(z)|^2 = (g(z)/N(z))^2|\rho'(z)|^2,$$

where z is height, but for convenience we will omit writing the altitude dependence, although in the analysis of lidar data it must of course be included.

Note that Equation (11) is also given by Hines (1960) in his Equation (36), as $\pm R/X = i(\gamma - 1)^{0.5}/C$, which is in effect the reciprocal of Equation (10) above. Hines (1960) notes that this equation is valid for the cases where $\omega << g/C$ - i.e. for waves of frequency much less than the Brunt-Vaisala frequency (made clearer by substituting $N/(\gamma - 1)^{0.5}$ (~ 1.6N) for g/C), with the additional constraint that $m^2 >> (\omega_a/C)^2$, i.e. small scale waves. The consequences of these assumptions will be considered in the next section.

4.5 Examination of the validity of the two assumptions used in deriving the equation $|U|^2 = (g/N)^2|\rho'|^2$ for monochromatic waves

4.5.1 The condition $m^2 >> w_a^2/C^2$

The minimum value of m allowed by the condition $m^2 >> (\omega_a/C)^2$ is around 2.2×10^{-4} rad m^{-1}, corresponding to a maximum vertical wavelength of $\lambda_z = 2\pi/m$ of around 29 km, assuming a maximum error of ~ 10% is acceptable. This is comparable to the maximum height range of the lidar data useful for gravity wave studies, which is from about 30 km altitude (i.e. above the aerosol contamination) to 60 to 70 km (where the falling signal-to-noise ratio of the data will limit the usefulness of the observations). Hence, the approximation $m^2 >> (\omega_a/C)^2$ would not appear to unduly constrain the usefulness of lidar data.

This approximation ($m^2 >> \omega_a^2/C^2$) is used to simplify the dispersion relation (compare Equations (3a) and 3(b)), by considering 'pure' gravity waves only. As the full acoustic-gravity wave dispersion relation was employed in the previous section in deriving Equation (11) it may be thought that the limit $m^2 >> \omega_a^2/C^2$ is overly restrictive, and that it should be possible to include much smaller m values into the analysis. However, Section 4.6 outlines some additional constraints on the allowable value of m, and the limit $m^2 >> \omega_a^2/C^2$ is in fact appropriate: for small scale waves (where $m^2 >> \omega_a^2/C^2$) the wave-induced fractional changes in density are equal in magnitude to the fractional changes in temperature,

which is critical for the interpretation of lidar data. Equation (11) is of course valid for small scale waves.

The approximation $m^2 >> \omega_a^2/C^2$ is also used in the next subsection (4.5.2).

The condition $m^2 >> \omega_a^2/C^2$ means that λ_z must be less than 29 km, as mentioned earlier. In the upper mesosphere, Senft *et al.* (1993) found the characteristic vertical wavenumber of the horizontal wind spectrum to be around 4.5×10^{-4} rad m^{-1}, equivalent to a vertical wavelength of around 14 km. That is, the wavenumber where the spectrum 'turns over' from a log-log slope of around -3 at high wavenumber, to a slope around zero for low wavenumber, is only about twice the minimum value allowed by Equation (10). Hall and Hoppe (1997) obtained a similar result, estimating a characteristic vertical wavenumber of around 4×10^{-4} rad m^{-1} (equivalent to a characteristic vertical wavelength of 16 km), for the height range of 70 to 90 km from EISCAT observations of vertical velocity. (Note, however, that Fritts (1995) suggests for the mesopause a characteristic vertical wavelength closer to 30 km, which is right at the upper bound for 'small scale waves'.) In any future analysis of lidar data vertical wave number spectra, derived from density or temperature fluctuations, the spectral slope of the 'low wave number range' should not be calculated using a lower bound that extends below the minimum wave number of 2.2×10^{-4} rad s^{-1} set by Equation (10). In practice this should always be the case, given the likely constraints on the useful altitude range of the lidar data.

4.5.2 The requirement that $\omega \leq 0.3N$

In this case, we require the ratio $|X|^2/|R|^2$ to be written in a form only in terms of known constants and ω, the wave frequency. This can be done by beginning with expressions for $|X|^2$ and $|R|^2$ as before, except now the k is replaced by ω/v, where v is the horizontal phase speed (see the definition of symbols above) and the dispersion relation is written as:

$$m^2 = (N^2 - \omega^2)/v^2 - (\omega_a^2 - \omega^2)/C^2,$$

to allow m to be substituted for. The full derivation is given in Appendix A.1, where we obtain the result:

$$|X|^2/|R|^2 = (N^2 - \omega^2)(g^2/N^4) \qquad (12)$$

under the constraint that $(v/C)^2 \sim 0$, that is, the horizontal phase speed is much less than the local speed of sound. In the middle atmosphere, this places a limit on the maximum magnitude of the horizontal phase speed of around 50 m s^{-1}. For $\omega^2 << N^2$, Equation (12) simplifies directly to yield the result quoted in Equation (11).

There are two points worth noting here. The first is that the constraint $(v/C)^2 \sim 0$ is essentially a restatement of the condition $m^2 >> \omega_a^2/C^2$. This follows directly from the definition of k, and the gravity wave dispersion relation, in the following manner. For $m^2 >> \omega_a^2/C^2$, the dispersion relation becomes $\omega^2 m^2 = (N^2 - \omega^2)k^2$. Thus substituting for m we have

$(N^2 - \omega^2)k^2/\omega^2 >> \omega_a^2/C^2$. As $k = \omega/v$, we require $(N^2 - \omega^2)/v^2 >> \omega_a^2/C^2$, or $v^2/(N^2 - \omega^2) << C^2/\omega_a^2$. Rewriting we see $v^2 << C^2(N^2 - \omega^2)/\omega_a^2$. The maximum value of the factor $(N^2 - \omega^2)/\omega_a^2$ is slightly less than unity in the middle atmosphere, for the case of $\omega << N$, as N is slightly less than ω_a. Therefore, the condition $m^2 >> \omega_a^2/C^2$ is equivalent to $v^2 << C^2$, or $(v/C)^2 \sim 0$.

Secondly, Equation (12) has been derived for the case which excludes the effects of rotation of the Earth's atmosphere on very low frequency waves. Including these effects modifies the expression to:

$$U'^2 = (g^2/N^4)[(\omega^2 + f^2)(N^2 - \omega^2)/(\omega^2 - f^2)]\rho'^2. \qquad (13)$$

This equation is derived in Appendix A.2, after the polarisation relation between horizontal and vertical velocity perturbations appropriate to a rotating atmosphere has been derived. (For $\omega^2 >> f^2$ the expression simplifies to Equation (12)). As Equation (13) is the more correct statement this will be used in the following.

Equation (13) shows how the ratio $|X|^2/|R|^2$ deviates from $(g/N)^2$ for different ω. This is illustrated in Figure 1, where the ratio $\alpha = \{|X|^2/|R|^2\}/(g/N)^2$ is plotted, as a function of ω. As can be seen, α falls off rapidly with increasing ω, from around 0.9 at $\omega = 0.3N$ (a 10% error, as expected), to around 0.2 for $\omega = 0.9N$, i.e. for a wave of $\omega = 0.9N$, the square of the horizontal wind fluctuation is only $0.2(g/N)^2$ times the normalised density fluctuation. Thus attempting to derive the horizontal velocity fluctuations due to gravity waves, from lidar observations of the normalised density fluctuations scaled only by $(g/N)^2$ as in Equation (11), will lead to results significantly too large for waves with frequencies greater than $\alpha = 0.3N$: the correct expression is given by Equation (13). Similarly, for very low frequency waves, the ratio α also deviates from unity, this time increasing with a steep gradient.

There is a simple physical interpretation of this result. For the gravity waves in the frequency range around $\omega \sim 0.05N$ to $\omega \sim 0.3N$ the wave motion is manifested in both horizontal and vertical velocity perturbations. As the wave frequency is increased, relatively more of the kinetic energy is present in the vertical velocity perturbation, W', and less (proportionally) is in U'. There is a natural link between ρ' and W', as the observed density fluctuations are due to vertical perturbation motion. Hence for the higher frequency waves, to convert the observed density changes back to horizontal motions requires multiplying ρ' by a factor less than unity - the quantity α as plotted in Figure 1. For the very lowest frequency waves (say $\omega < 0.05N$) there is little vertical motion, and hence little density change. To scale the observed density changes back to horizontal velocity thus requires multiplication by a number significantly greater than unity.

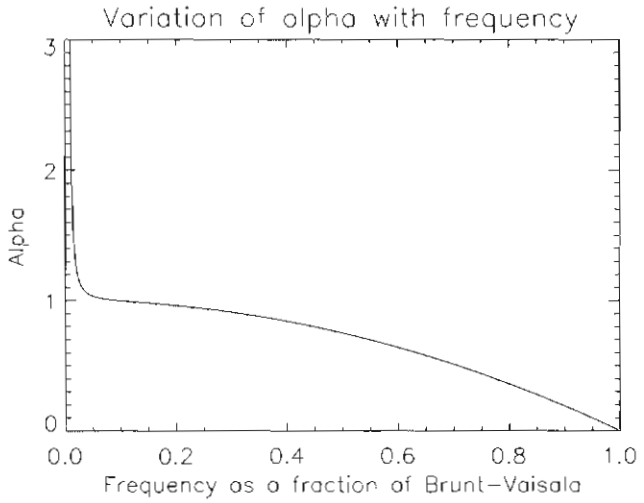


Figure 1. *Correction factor, α, versus frequency (plotted as a fraction of the Brunt-Vaisala frequency, N) for obtaining wave-induced horizontal wind variance from the square of the wave-induced fractional density fluctuations multiplied by $(g/N)^2$. At high-frequencies scaling the density fluctuation by $(g/N)^2$ leads to an overestimate of the horizontal wind variance, hence the multiplicative correction factor, α, is less than one. At very low frequencies, where there is very little vertical velocity perturbation and hence little density change, the correction factor becomes very large.*

For high frequency waves ($\omega > 0.3N$) it is not strictly appropriate to compare measured density fluctuations with horizontal wind fluctuations. However, the observed gravity wave ω-spectra suggests that waves with $\omega > 0.3N$ are a relatively minor component of the total wave population (i.e. $F_{GW}(\omega) \propto \omega^{-5/3}$). That is, when comparing vertical wave number spectra of density changes with vertical wave number spectra of horizontal wind changes, where both of these spectra include the fluctuations from waves of all frequencies, the relative contribution from the high frequency waves to the vertical wave number spectra is essentially negligible. For the lowest frequency waves it is also not valid to scale density fluctuations to horizontal velocity perturbations. In general however, due to the limited duration of most observing intervals, it is possible to ignore the effects of such low frequency waves by averaging out the long term variations in the data.

These points will be considered further in Section 4.8, where it will also be shown it is valid to derive 'potential energy per unit mass' from density or temperature fluctuations, for waves of all frequencies,

using $(g/N)^2$ scaling alone. Section 4.7 considers the relationship between energy per unit mass and relative density fluctuations. Before considering the energy per unit mass however, one further point will be introduced, that of the conditions necessary for the equivalence of the wave induced density and temperature fluctuations.

4.6 Conditions required for the equivalence of wave-induced density and temperature fluctuations

It is often noted in the lidar literature that the wave induced normalised fluctuations in temperature, T′, are equivalent to the wave induced normalised fluctuations in density, ρ′, at least in the sense that the magnitudes are equal, i.e. $(T')^2 = (\rho')^2$. This is true only under certain conditions, that of 'pure gravity waves', where the only wave induced density changes of the atmosphere are from buoyancy forces. This section will consider this aspect of lidar data interpretation. Much of this section, and a significant part of the next, comes from the very clear and comprehensive work of Makhlouf *et al.* (1990), which is recommended for further reading. Note though that Makhlouf *et al.* (1990) do not consider the effects of the Earth's rotation, by their choice of the 'non-rotating' dispersion and polarisation equations (Equations (1) and (5) respectively above). We return to this point later in this paper.

For an ideal gas, the change in temperature, due to changes in pressure and density, can be written as

$$\Delta T/T_0 = \Delta p/p_0 - \Delta\rho/\rho_0, \qquad (14)$$

where the subscript indicates unperturbed quantities. Makhlouf *et al.* (1990) obtain this result from the linearised ideal gas law. The validity of this equation can be confirmed by considering the gas law: $p = nkT$, where n is the number density and k is Boltzmann's constant. Rewriting this gives $p = \rho(k/m_i)T$, where m_i is the mass of an individual atom or molecule of the gas, and $\rho = nm_i$. The term k/m_i is approximately constant in the middle atmosphere. (In fact, as the mass of an individual gas atom or molecule is given by the molecular weight, M_W, divided by Avagadro's number, and as R, the universal gas constant equals $N_A k$, we can write $k/m_i = R/M_W$.) Hence $T \propto (p/\rho)$, so that small changes in T are proportional to changes in p, but opposite to changes in ρ.

Makhlouf *et al.* (1990) also quote a result for Δp, the change in pressure, derived from the linearised horizontal momentum conservation equation, for monochromatic waves, of:

$$\Delta p = vU'\rho_0, \qquad (15)$$

where v is the horizontal phase speed, and U′ is the horizontal perturbation velocity, as before. The mean pressure and density are also related through the equation

$$C^2 = \gamma p_0/\rho_0. \qquad (16)$$

Hence, substituting for ρ_0 in (15) using (16) gives $\Delta p = vU'\gamma p_0/C^2$, so $\Delta\rho/p_0 = \gamma U'(v/C^2)$, hence Equation (14) becomes

$$\Delta T/T_0 = \gamma U'(v/C^2) - \Delta\rho/\rho_0. \quad (17)$$

The fractional density change, $\Delta\rho/\rho_0$, induced by a gravity wave is composed of two components, the change due to the vertical displacement of the air parcel into a region of differing pressure, denoted $\Delta\rho_a/\rho_0$ by Makhlouf *et al.* (1990), and the density change induced by the acoustic compression of the wave itself, $\Delta\rho_c/\rho_0$, i.e.

$$\Delta\rho/\rho_0 = \Delta\rho_a/\rho_0 + \Delta\rho_c/\rho_0.$$

Makhlouf *et al.* (1990) demonstrate that $\Delta\rho_c/\rho_0 = U'v/C^2$, hence Equation (17) becomes:

$$\Delta T/T_0 = (\gamma - 1)U'(v/C^2) - \Delta\rho_a/\rho_0. \quad (18)$$

For the case of small vertical scale waves, where $v^2 << C^2$, the first term of Equation (18) can be neglected. This approximation was shown above to be the same as requiring $m^2 >> \omega_a^2/C^2$. That is, the approximation of $v^2 << C^2$, which is equivalent to requiring 'small' vertical wavelengths, is also equivalent to the requirement that the waves are 'pure' gravity waves, with no acoustic (pressure) component. Hence temperature changes that would be brought about by pressure changes from the wave itself can be ignored. For these small scale, 'pure' gravity waves then, we have:

$$\Delta T/T_0 = - \Delta\rho_a/\rho_0. \quad (19)$$

So the wave induced fractional temperature fluctuations are equal and opposite to the fractional density fluctuations in this case.

As was discussed earlier, the requirement that $m^2 >> \omega_a^2/C^2$ will nearly always be satisfied for the waves observed with the lidar, so that Equation (19) will also be generally satisfied.

Note too that the requirement of small scale waves also appears to be a restatement of the often quoted Boussinesq approximation, e.g. Salby (1996): "The Boussinesq approximation neglects compression of an air parcel introduced by changes of pressure, retaining only density variations that are introduced thermally".

4.7 Relationship between wave energy per unit mass and normalised density and temperature fluctuations

Following Makhlouf *et al.* (1990), the time averaged kinetic energy per unit volume is:

$$\langle KE_V \rangle = {}^1/_4\rho_0(|U'|^2 + |W'|^2) \quad (20)$$

where U' and W' were defined above, and the factor of ${}^1/_4$ appears as this is a time average, hence we take U'_{RMS} (= ${}^1/_{\sqrt{2}}U'$) and W'_{RMS} (= ${}^1/_{\sqrt{2}}W'$),

which gives the additional factor of $^1/_2$ in the equation. Note that Makhlouf *et al.* (1990) use the symbol KE for the time averaged kinetic energy per unit volume, while here $<KE_V>$ is used to make the meaning more apparent.

Makhlouf *et al.* (1990) show that this equation can be written as:

$$<KE_V> = {}^1/_4\rho_0 N^2|h|^2 [(1 - (\omega/N)^2(v/C)^2)/(1 - (v/C)^2)] \quad (21)$$

where h is the vertical displacement due to the wave, and the other symbols are as before. This was done by using the dispersion relation, which relates *m* and *k*, to express U′ in terms of W′. For the small scale waves (pure gravity waves) $v^2/C^2 \sim 0$, so Equation (21) becomes

$$<KE_V> = {}^1/_4\rho_0 N^2|h|^2 \quad (22)$$

Makhlouf *et al.* (1990) also derive, again for 'small scale' waves where compressional temperature changes can be ignored, the relation:

$$\Delta T/T_0 = (\gamma - 1)|h|/(\gamma H). \quad (23)$$

As $H = C^2/\gamma g$ (from $H = kT/mg$, $p_0 = nkT$, $\rho_0 = \gamma p_0/C^2$, $\rho_0 = nm$), it is straightforward to show that $<KE_V> = {}^1/_4\rho_0 C^2/(\gamma-1)(\Delta T/T_0)^2$, and hence:

$$<KE_V> = {}^1/_4\rho_0(g/N)^2(\Delta T/T_0)^2. \quad (24)$$

The time averaged kinetic energy per unit mass, $<KE_M>$, is simply $<KE_V>/\rho_0$, so that:

$$<KE_M> = {}^1/_4(g/N)^2(\Delta T/T_0)^2, \quad (25)$$

and as $\Delta T/T_0 = -\Delta\rho_a/\rho_0$ (Equation (19))

$$<KE_M> = {}^1/_4(g/N)^2(\Delta\rho/\rho_0)^2. \quad (26)$$

Note the validity of Equations (25) and (26) is not limited to $\omega \leq 0.3N$. Equation (26) is of course similar to Equation (11) above, viz.: $|U'|^2 = (g/N)^2|\rho'|^2$ which relates the wave induced changes in density to the horizontal velocity perturbations, apart from the factor of $^1/_4$ from the time averaging, and with the important difference that Equation (11) is not valid for all $\omega > 0.3N$. (See Section 4.5.2.) The similarity between Equations (11) and (26) is physically justifiable - for $\omega \leq 0.3N$ or so, most of the wave velocity perturbation (and hence kinetic energy) is present in the horizontal velocity – so the time averaged kinetic energy per unit mass is to a good approximation equal to $^1/_4|U|^2$. At higher frequencies, proportionally more energy is present in the vertical velocity perturbation, hence $^1/_4|U|^2$ is no longer a good measure of $<KE_M>$.

There is a low frequency limit to the generality of Equations (21)–(26). This arises because at wave frequencies comparable to the inertial frequency, *f*, the dispersion and polarisation equations used in the deriving the above equations are no longer correct. The appropriate dispersion relation which takes 'rotation' into account can be written as

$$m^2 = k_H^2(N^2 - \omega^2)/(\omega^2 - f^2), \qquad (27)$$

from Equation 8.4.13 of Gill (1982, p.258), where $k_H^2 = k^2 + l^2$ (= total horizontal wavenumber), while the polarisation equation relating U'^2 and W'^2 is now (from Gill, 1982, derived from equations given on p.263)

$$U'^2 = [(\omega^2 + f^2)(N^2 - \omega^2)]/[(\omega^2 - f^2)\omega^2]W'^2. \qquad (28)$$

The derivation of Equation (28) is given in Appendix A.2.

For $\omega \geq 0.3f$ (corresponding to periods less than about 6 h at Kingston, and periods less than about 4 h at Davis) Equations (25) and (26) will be at most in error by 10%, decreasing in error with increasing frequency. Long period waves would not be fully resolvable in an observing interval of say 8 to 12 hours. There is however significant power at the low frequencies, from both long period gravity waves and also from tidal motions. In general, the effects of these low frequency motions can be filtered out by averaging in the time domain. Hence for studies of intermediate to short period gravity waves the dispersion relation given as Equation (1) can be used, but it is important to be aware of the limitations. The effects of very low frequency waves (comparable to the inertial frequency) are further considered below, but here we state the result (derived as Equation (34) below) that when the effects of rotation are taken into account, Equation (26) becomes:

$$<KE_M> = {}^1/_4(g/N)^2[(\omega^2 + f^2)/(\omega^2 - f^2)]\rho'^2.$$

For $\omega >> f$ this equation simplifies to Equation (26).

The time averaged potential energy per unit mass is given by:

$$<PE_M> = {}^1/_4(g/N)^2\rho'^2.$$

This equation can be derived from Equation 6.7.6 of Gill (1982) and is of course identical with Equation (26) above, with the important difference that this is valid for all wave frequencies. Hence for complete generality it is more correct to interpret the ${}^1/_4(g/N)^2\rho'^2$ as potential energy per unit mass rather than kinetic energy per unit mass. The differences are only of importance at very low frequencies, and in practice will normally be negligible. A further discussion of the effects of low frequency waves is given in the following sections. Note that it is the equivalence of kinetic and potential energies that allows one to relate the density variations (linked directly to vertical motions and hence also to potential energy changes) to the horizontal velocity variations (linked to the kinetic energy). This point will be amplified below.

In summary, a simple $(g/N)^2$ scaling of the square of the wave-induced fractional temperature or density fluctuations gives a valid measure of the kinetic energy per unit mass for a monochromatic wave, except for very low frequencies. At intermediate frequencies, $3f \leq \omega \leq 0.3N$, this scaling also leads to an estimate of the horizontal velocity perturbations, with a maximum error around 10% at the frequency limits. Comparisons

of gravity wave spectra from lidar and other instruments (e.g. radar, sonde flight data) can hence be obtained from such scaling, but some caution is required in interpreting the results. The consequences of the frequency dependence of some of the above results will be explored in the next section.

4.8 Derivation of gravity wave spectra

4.8.1 Horizontal wind and relative temperature (or density) spectra

The relationships derived above for monochromatic waves can be used in the derivation of horizontal wind variance spectra from lidar observations of relative temperature fluctuations, but, as will be shown, a simple scaling using these relationships is in general valid only for waves of mid-range frequencies ($3f \leq \omega \leq N/3$).

We use the expression relating vertical and horizontal wind fluctuations (Equation 28) to write:

$$W'^2 = U'^2\omega^2(\omega^2 - f^2)/[(N^2 - \omega^2)(\omega^2 + f^2)]. \quad (29)$$

From Equation (23) above (using $\rho' = T'$) and $h = W'/i\omega$, we have:

$$\rho'^2 = W'^2N^4/(\omega^2g^2), \quad (30)$$

and substituting for W'^2 gives:

$$\rho'^2 = U'^2[N^4/g^2](\omega^2 - f^2)/[(N^2 - \omega^2)(\omega^2 + f^2)]. \quad (31)$$

For $f << \omega << N$ this equation simplifies directly to Equation (11).

We can write Equation (31) in terms of gravity wave spectra as:

$$F_\rho(m, \omega) = F_U(m, \omega)[N^4/g^2](\omega^2 - f^2)/[(N^2 - \omega^2)(\omega^2 + f^2)].$$

It is usual to assume both the density and horizontal wind spectra are separable in vertical wave number and frequency, e.g. $F_U(m, \omega) = F_U(m)B_0\omega^{-p}$ (Senft *et al.*, 1993), where B_0 is

$$B_0 = (p - 1)f^{p-1}/[1 - F^{p-1}],$$

with $F = f/N$, and $p = 5/3$. Note that $\int_f^N B_0\omega^{-p}d\omega = 1$. Hence, integrating w.r.t. ω gives:

$$F_\rho(m)\int_L^T B_0\omega^{-p}d\omega = N^4/g^2 \int_L^T B_0(\omega^2 - f^2)/[(N^2 - \omega^2)(\omega^2 + f^2)]\omega^{-p}d\omega F_U(m) \quad (32)$$

where T and L are the upper and lower limits of the integration respectively. The integration can be done in a straightforward manner by numerical methods (using the inbuilt functions in the Interactive Data Language software). Note that the function is not defined at $\omega = N$. The upper limit of the integral is therefore treated as 'open'.

Transposing the terms, this can be written as

$$F_U(m) = \Pi g^2/N^2 F_\rho(m) \qquad (33)$$

where Π is the ratio of the definite integrals as given in Equation (32). The value of Π depends on the limits chosen for the integration. Some representative values are given in Table 1.

Table 1. *Representative values of Π, the ratio of the definite integrals as given in Equation (32).*

Integration range (ω)	Π	Notes
$f \leq \omega < N$	1.37	All frequencies except $\omega = N$
$f \leq \omega \leq 3f$	2.41	low frequencies
$3f \leq \omega \leq N/3$	1.05	mid frequencies
$N/3 \leq \omega < N$	0.33	high frequencies
$10f \leq \omega \leq N/2$	0.97	Typical observing bandwidth (0.3-day)

From Table 1, it is clear that for an integration range in ω (equal to a bandwidth in frequency) taken in the mid-frequency range then Equation (32) is effectively equivalent to Equation (11), which was derived for monochromatic waves, as Π is within a few percent of unity. For an integration range that includes very high frequency waves however, the factor Π differs significantly from unity. This is because at the high frequencies there is little horizontal motion and hence the observed density variation (from vertical motion as noted earlier) must be therefore scaled by a factor less than $(g/N)^2$ in order to estimate the horizontal velocity changes. At very low frequencies observing the density variations will underestimate the horizontal velocity, hence the factor is greater than unity. Setting the integration limits determines the contribution of the different frequency waves that are included in the summation (from the $\omega^{-5/3}$ spectral shape), and hence this determines the value of Π.

4.8.2 Relative density and kinetic energy per unit mass spectra

Using Equation (29) to write the horizontal wind fluctuations in terms of the vertical wind fluctuations, and substituting into Equation (20) for the kinetic energy per unit volume, we have:

$$\langle KE_M \rangle = {}^1/_4 W'^2[1 + (N^2 - \omega^2)(\omega^2 + f^2)/\omega^2(\omega^2 - f^2)],$$

where $\langle KE_M \rangle = \langle KE_V \rangle/\rho_0$ as before. As $W'^2 = (g^2/N^4)\omega^2\rho'^2$ (from Equation (30)), we obtain:

$$\langle KE_M \rangle = {}^1/_4(g^2/N^4)\rho'^2(N^2\omega^2 + N^2f^2 - 2\omega^2f^2)/(\omega^2 - f^2),$$

and on taking out a common factor of N^2 from the numerator, gives:

$$<KE_M> = {}^1/_4(g^2/N^2)\rho'^2(\omega^2 + f^2 - 2\omega^2 f^2/N^2)/(\omega^2 - f^2).$$

For $\omega \sim f$, the term $2\omega^2 f^2/N^2$ is $\sim 2f^4/N^2$, which is much less than f^2. For $\omega \sim N$, this mixed term becomes $\sim 2f^2$, which is much less than N^2. Hence ω^2 is always much bigger than $2\omega^2 f^2/N^2$, so this mixed term can be neglected. Hence we have:

$$<KE_M> = {}^1/_4(g^2/N^2)\rho'^2(\omega^2 + f^2)/(\omega^2 - f^2). \quad (34)$$

On transposing:

$$\rho'^2 = 4<KE_M>(N^2/g^2)(\omega^2 - f^2)/(\omega^2 + f^2).$$

Changing into the spectral form, as was done for Equation (31), yields:

$$F_\rho(m)\int_L^T B_0\omega^{-p}d\omega = 4F_{KEM}(m)(N^2/g^2)\int_I^T B_0(\omega^2 - f^2)/(\omega^2 + f^2)\omega^{-p}d\omega. \quad (35)$$

This can be written as:

$$F_{KEM}(m) = {}^1/_4\Theta(g/N)^2 F_\rho(m) \quad (36)$$

where Θ is the ratio of the definite integrals as given in Equation (35). The value of Θ for a selection of integration limits is given in Table 2.

Table 2. *Representative values of Θ, the ratio of the definite integrals in Equation (35).*

Integration range (ω)	Θ	Notes
$f \le \omega \le N$	1.51	all frequencies
$f \le \omega \le 3f$	2.41	low frequencies
$3f \le \omega \le {}^N/_3$	1.06	mid frequencies
${}^N/_3 \le \omega < N$	1.00	high frequencies
$10f \le \omega \le {}^N/2$	1.00	typical observing bandwidth (0.3-day)

Similar comments apply here as for the value of Π given above. Note that in this case the value of Θ quickly converges to unity as the lower limit of integration, L, increases. It is clear that if the lower frequency ranges are excluded (i.e. excluding those frequency ranges where the wave induced horizontal velocities are relatively large) then there is a direct equivalence between density changes (intrinsically linked to vertical motions) and the kinetic energy. It is also clear that interpreting the $(g/N)^2$ scaled density spectrum as kinetic energy per unit mass

spectrum has applicability over a greater range of wave frequencies than interpreting this quantity as an estimate of the horizontal wind fluctuation spectrum.

4.8.3 Relative density and 'available' potential energy per unit mass spectra

Gill (1982, Section 6.7) gives an expression relating the 'available' potential energy per unit density and the square of the normalised density fluctuations. This can be written as the time averaged available potential energy per unit mass as:

$$<PE_M> = {}^1/_4(g/N)^2\rho'^2.$$

As there is no explicit ω-dependence in the relationship, in spectral form this is simply:

$$F_{PEM}(m) = {}^1/_4(g/N)^2F_\rho(m). \qquad (37)$$

Hence scaling the density spectra by $(g/N)^2$ is valid for all frequencies in this case. Clearly, for maximum validity, interpreting the scaled density spectra as available potential energy per unit mass spectra is to be preferred, as the density fluctuations are intrinsically linked to the vertical motions (and hence to potential energy changes). As shown by Table 2 however, interpreting the scaled density spectra as kinetic energy per unit mass spectra is also valid when excluding waves of very low frequency.

In the light of the above discussion of kinetic and potential energies per unit mass, it is worth noting once more that the relationship between horizontal winds and the density fluctuations is in reality a statement of the equivalence of the potential and kinetic energies. The density fluctuations are intrinsically related to the potential energy changes, as both result from vertical motion of air parcels due to the waves. Except at very low frequencies, comparable to the inertial frequency, the potential energy (Equation (37)) and kinetic energies (Equation (34)) are equal. Excluding the highest frequencies ($\omega \sim N$), the wave kinetic energy is dominated by the horizontal velocity perturbations – hence for the 'intermediate frequency range' we can link the horizontal velocities to the density variations via the kinetic and potential energies. Physically however, the horizontal velocity perturbations do not give rise to density perturbations (under the usual assumptions of 'pure' gravity waves in an incompressible atmosphere), hence linking the two, while mathematically possible, tends to confuse the physical processes in operation.

4.8.4 Comparison with the formalism of Fritts and VanZandt (1993)

For completeness, a comparison of the above with the formalism of Fritts and VanZandt (1993) will be given.

Fritts and VanZandt (1993) assume a total energy per unit mass spectrum of the form:

$$E_M(\mu, \omega, \phi) = E_0 A(\mu) B(\omega) \Phi(\phi),$$

where μ is the normalised wavenumber, $\mu = m/m_*$, where m_* is the characteristic wavenumber discussed earlier, $B(\omega)$ is as before, and $\Phi(\phi)$ represents the azimuthal dependence. (The subscripted upper case M denotes a 'per unit mass' quantity, as before.) Fritts and VanZandt (1993) show in their Equation (7) that the variance spectrum of the vertical wind perturbation, E_W (i.e. the spectrum of W', the vertical velocity perturbations), is given by:

$$E_W(\mu, \omega, \phi) = (\omega/N)^2 E_M(\mu, \omega, \phi)\delta\_(\omega),$$

where $\delta\_(\omega) = 1 - (f/\omega)^2$. Fritts and VanZandt (1993) state that this has been derived using the wave polarisation equations, which allow for non-hydrostatic ($\omega \sim N$) and rotational (i.e. inertial) effects ($\omega \sim f$).

As $W'^2 = (g^2/N^4)\omega^2\rho'^2$ (from Equation (30) above), we can replace the vertical wind variance spectrum by the density variance spectrum, and find:

$$E_\rho(\mu, \omega, \phi)(g^2/N^4)\omega^2 = (\omega/N)^2 E_M((\mu, \omega, \phi)\delta\_(\omega),$$

or on transposing and simplifying, we find for the total energy per unit mass:

$$E_M(\mu, \omega, \phi) = [1/\delta\_(\omega)](g^2/N^2)E_\rho(\mu, \omega, \phi),$$

where $E_\rho(\mu, \omega, \phi)$ is the variance spectrum of the density fluctuations, ρ'. (i.e. $E_\rho(\mu, \omega, \phi)$ is the spectrum of ρ'.) The similarity between this equation and Equation (37) above, relating potential energy per unit mass and the square of the density fluctuations, is apparent.

The term $[1/\delta\_(\omega)]$ can be rewritten as $\omega^2/(\omega^2 - f^2)$. Hence, following the formalism of Fritts and VanZandt (1993), we conclude that the spectrum of the total energy per unit mass spectrum of the wave fluctuations can be derived from the variance spectrum of the density fluctuations, using the equation:

$$E_M(\mu, \omega, \phi) = [\omega^2/(\omega^2 - f^2)](g^2/N^2)E_\rho(\mu, \omega, \phi)).$$

As noted, this is similar to Equation (37), relating the time averaged potential energy per unit mass spectrum to the density variance spectrum. The differences arise as one is a non-averaged total energy, while the other is the time averaged potential energy. The consistency between Fritts and VanZandt (1993) and the equations derived above (from Gill, 1982) can easily be shown, as follows. The maximum kinetic and potential energies are simply twice the time averaged values, hence from Equations (34) and (37) we can write, in full (m, ω, ϕ) spectral form:

$$KE_{M(max)}(m, \omega, \phi) = 0.5(g/N)^2 E_\rho(m, \omega, \phi)(\omega^2 + f^2)/(\omega^2 - f^2), \text{ and}$$

$$PE_{M(max)}(m, \omega, \phi) = 0.5(g/N)^2 E_\rho(m, \omega, \phi).$$

Adding together, to get the maximum *total* energy, we obtain:

$$E_{M(max)}(m, \omega, \phi) = 0.5(g/N)^2 E_\rho(m, \omega, \phi)(2\omega^2)/(\omega^2 - f^2),$$

or

$$E_{M(max)}(m, \omega, \phi) = (\omega^2)/(\omega^2 - f^2)(g/N)^2 E_\rho(m, \omega, \phi).$$

This equation is the exact equivalent of that derived above relating total energy per unit mass to the variance of the density fluctuations obtained from the formalism of Fritts and VanZandt (1993).

4.9 Discussion and comparison with published analyses

With the above derivations having established the basis for the interpretation of lidar data, this section makes a brief survey of some recent published accounts of gravity wave analysis. This is not a comprehensive literature review, just a selection to indicate how lidar data have been used. The foregoing material allows the relationships between the different methods to be seen.

The estimate of the horizontal wind fluctuations from observed density (or temperature) fluctuations (Equation (11)) has often been used to derive gravity wave spectra from lidar data. For example, $(g/N)^2$ scaling of the density variations has been used by Gardner *et al.* (1989), their Equation (11), given in the form $|U'|^2 = (\gamma HN/(\gamma - 1))^2|\rho'|^2$, is equivalent to $(g/N)^2$ scaling, as $H = C^2/\gamma g$); Senft and Gardner (1991); Senft *et al.* (1993); Hostetler and Gardner (1994); Meriwether *et al.* (1994); and Collins *et al.* (1996). As noted, this approach is valid to within a few percent or so for $0.3N \geq \omega \geq 3f$ (Table 1). As also noted, the equivalence between the density and horizontal wind variances is a consequence of the equivalence of potential and kinetic energies.

Mitchell *et al.* (1991) relate the fractional density changes to the potential energy per unit mass of the fluctuation, using the relation, quoted as coming from Gossard and Hooke (1975):

$$PE_M = {}^1/_2(g/N)^2|\rho'|^2, \tag{38}$$

with a note that this is valid for all frequencies (as discussed above). Mitchell *et al.* (1991) also quote the result, again with a reference to Gossard and Hooke (1975), relating the time averaged wave kinetic and potential energies as:

$$\langle PE\rangle/\langle KE\rangle = (\omega^2 - f^2)/(\omega^2 + f^2), \tag{39}$$

where f is the inertial frequency as before. For $\omega >> f$, this ratio is unity. Equation (38) is not a time averaged quantity, but is in fact the maximum potential energy of the wave, as indicated by the factor of ${}^1/_2$ at the front. However, from Equations (26) and (39), it is also clear that the maximum PE_M must also be twice the time averaged KE_M, for $\omega >> f$. Equations (38) and (39) are clearly consistent with Equation (34) derived above, for the time averaged kinetic energy per unit mass (with rotation). As noted earlier Equation (38) is valid for all frequencies, as the potential energy is directly related to the density changes through the vertical displacement.

While Gossard and Hooke (1975) was quoted as the source of Equations (38) and (39), a search through this reference did not locate these equations. However, Gill (1982) derives an equation very similar to Equation (38) above (his Equation 6.7.6), and also derives the ratio of kinetic energy to potential energy (Equation 8.6.6) as:

$$KE/PE = (\omega^2 + f^2\sin\phi')/(\omega^2 - f^2\sin\phi') \qquad \text{(Gill, 1982, eqn. 8.6.6)}$$

where ϕ' is defined by:

$$\omega^2 = f^2\sin^2\phi' + N^2\cos^2\phi'. \qquad \text{(Gill, 1982, eqn. 8.4.15)}$$

Physically, the angle ϕ' is the angle of energy propagation of the wave, relative to the horizontal. While at first look this equation from Gill (1982) and Equation (39) appear inconsistent, it is clear that for $\phi' = \pi/2$ the expressions are equivalent, as at $\phi' = \pi/2$, ω takes the value of f, from Equation 8.4.15 of Gill. As ϕ' decreases from this maximum value, ω increases. At $\phi' = \pi/4$, $\omega^2 = {}^1/_2(N^2 + f^2) \approx N^2/2 >> f^2$, so that functionally, Equation 8.6.6 of Gill and Equation (39) are similar. For completeness, this point is considered in more detail in Appendix A.3, where the mathematical equivalence is demonstrated, and a brief comment is made as to why the ratio KE/PE differs from unity for waves of very low frequency.

Mitchell *et al.* (1991) go on to note that they prefer to take the spectrum of a quantity K, which they define as

$$K = (g/N)\rho'. \qquad (40)$$

In a later paper, Mitchell *et al.* (1996) give the analogue of Equation (13), but use the dispersion relation valid for all frequencies (Equation (27) above) and quote:

$$U' = \sqrt{[(1 - \omega^2/N^2)/(1 - f^2/\omega^2)]}(g/N)\rho', \qquad (41)$$

noting that for $N >> \omega >> f$ this allows (g/N) scaling from ρ' or T' to the horizontal wind fluctuations, as has been discussed earlier. On squaring, this expression can be written as:

$$U'^2(\omega^2 - f^2) = \omega^2(N^2 - \omega^2)(g^2/N^4)\rho'^2.$$

When comparing with Equation (31) derived above, it is seen that these equations are not consistent, as Equation (31) written in this form is:

$$U'^2(\omega^2 - f^2) = (\omega^2 + f^2)(N^2 - \omega^2)(g^2/N^4)\rho'^2.$$

The difference appears to be that the equation quoted by Mitchell *et al.* (1996) has been derived using the 'non-rotation' polarisation relationship for U' and W', and hence is not valid at very low frequencies. Mitchell *et al.* (1996) however do not make use of this incorrect equation in their analysis, but again calculate potential energy per uint mass spectra, once more using Equation (38) which is a valid interpretation of their lidar data.

Whiteway and Carswell (1994) quote a result for the 'average available potential energy per unit mass' (from Gill, 1982) as:

$$E_p = {}^1/_2(g/N)^2\langle(T')^2\rangle, \qquad (42)$$

which they use in their analysis of lidar data. As the factor at the front of the right hand side is $^1/_2$, and not $^1/_4$ as would be expected for a 'time average', it might be thought that there is an inconsistency present. However, it appears they define their T' as the root mean square of the temperature fluctuations, which accounts for the difference. Hence this is also a valid approach, and is effectively identical with the Mitchell *et al.* (1996) method.

In practice, the analysis performed on the lidar data will depend on the problem under investigation, for example, whether a comparison with data sets obtained by other instruments is to be performed, or whether a study of the energy density spectrum is intended. Comparisons with horizontal velocity fluctuations will be subject to the uncertainty as to the frequency range of the waves contributing to the study, and hence to the applicability of Equation (11), although the $-(^5/_3)$ (log-log) slope of the ω-spectrum reduces this effect as most wave power is present at frequencies much lower than $\omega = N$. In any case, interpreting the $(g/N)^2$ scaled square of the wave-induced density and/or temperature relative fluctuations as an 'energy per unit mass' quantity appears to be preferable. Interpreting the scaled squared density (or temperature) as a potential energy per unit mass has the advantage of being valid for all frequencies. Equally, for lidar–lidar comparisons, simply calculating density or temperature spectra is sufficient.

Senft *et al.* (1993) in their presentation of lidar data derive horizontal velocity spectra from observed density fluctuations, using a relation between these quantities of the form:

$$F_U'(m) = {}^1/_b(g/N)^2F_\rho(m). \qquad \text{(Senft } et\ al.\text{, 1993, eqn. 7)}$$

A value of $b = {}^3/_4$ is obtained from an integration performed over all frequencies ($f \rightarrow N$). However they use a value of $b = 1$, by noting that their observations do not extend to the very lowest frequency waves.

Senft *et al.* (1993) quote the polarisation relation they use as:

$$W'^2 = (k/m)^2U'^2 \qquad \text{(Senft } et\ al.\text{, 1993, eqn. 3)}$$

together with the dispersion relation:

$$m^2 = k^2N^2/(\omega^2 - f^2). \qquad \text{(Senft } et\ al.\text{, 1993, eqn. 4)}$$

This dispersion relation is not valid for very high frequency waves ($\omega \sim N$), while the polarisation relation used is only valid for mid-frequency waves. (The more general expressions were given in the previous section.) However their integration is performed over a large range of frequencies when deriving their results. Their spectral equation is:

$$F_\rho(m) \int_L^T B_0 \omega^{-p} d\omega = F_U(m)(N/g)^2 \int_1^T B_0 (1 - f^2/\omega^2)\omega^{-p} d\omega.$$

This can be solved analytically. The integral on the right hand side becomes:

$$RHI = B_0 \left| \omega^{1-p}/(1 - p) + f^2 \omega^{-1-p}/(1 + p) \right|_L^T$$

Evaluating for all frequencies, from $\omega = f$ to $\omega = N$, yields a value of RHI = 0.7425 to four decimal places, while that on the left hand side is unity in this case. The result of $b = {}^3/_4$ (= 0.75) obtained by Senft *et al.* (1993) follows if the approximation $(f/N) = 0$ is made in the expression for B_0.

Table 3 compares the ratios of the definite integrals for scaling density spectra to horizontal velocity variance spectra derived using the full dispersion and polarisation relations (as was done in Section 4.8.2) and the approximate relations (as used by Senft *et al.*, 1993). Data for the 'full' relations are reproduced from Table 1.

Table 3. *Comparison of ratios of the definite integrals for scaling between density and horizontal wind spectra when using the 'full' and 'approximate' dispersion and polarisation relations.*

Integration range (ω)	'Full'	'Approximate'
$f \leq \omega < N$	1.37	1.35
$f \leq \omega \leq 3f$	2.41	1.84
$3f \leq \omega \leq {}^N/_3$	1.05	1.03
${}^N/_3 \leq \omega < N$	0.33	1.00
$10f \leq \omega \leq {}^N/2$	0.97	1.00

As can be seen from Table 3, for mid-range frequencies, the results are comparable. The differences are greatest when comparing the lowest and highest ranges for each case. This is because the full effects of rotation are not included in the case of the lowest frequencies in the polarisation relation used by Senft *et al.* (1993), and because the dispersion relation they used is not valid for $\omega \sim N$ for the highest frequencies. In the formalism used by Senft *et al.* (1993), the density variance due to a high frequency wave would be scaled by $(g/N)^2$ to give the horizontal wind variance, whereas in fact there is little horizontal velocity perturbation associated with the highest frequency gravity waves. While high frequency gravity waves are a relatively minor component of the observed gravity wave spectrum, this point should be borne in mind if it is required to derive horizontal wind variance spectra from Rayleigh lidar data.

Allen and Vincent (1995) investigated gravity wave properties using temperature fluctuations from balloon borne radiosonde data. They gave an expression relating the energy density, E_0, to the total normalised temperature variance:

$$E_0 = (g/N)^2(T')^2/(B_0C_{In}), \quad (43)$$

where

$$B_0 = (p-1)\, f^{p-1}/[1 - F^{p-1}],$$

as above, and

$$C_{in} = f^{1-p}/(1 - F^2)[(1 - F^{p-1})/(p-1) - (1 - F^{p+1})/(p+1)] \quad (44)$$

with $F = f/N$, and $p = {}^5/_3$, as before. (B_0 comes from Fritts and VanZandt, 1993; C_{In} comes from integrating their Equation (5) - see below - from $\omega = f$ to $\omega = N$, i.e. over *all* frequencies.)

Making the approximation $F = 0$ (good to a few percent), then the product B_0C_{In} becomes:

$$B_0C_{In} = (p-1)f^{p-1}\, f^{1-p}\, [(p+1) - (p-1)]/[(p+1)(p-1)] = {}^3/_4, \quad (45)$$

as $p = {}^5/_3$. Hence, Equation (43) is, approximately,

$$E_0 = {}^4/_3(g/N)^2\, (T')^2. \quad (46)$$

That is, the multiplicative constant in the scaling equation is identical with that derived above for the 'approximate' equations. (see Table 3).

The spectral equation used by Allen and Vincent (1995), their Equation (5), in terms of the symbols used here, is:

$$F_T = (N/g)^2[(1 - f^2/\omega^2)/(1 - f^2/N^2)]F_E,$$

where F_E is the energy spectrum. Note that apart from the (constant) denominator of $1 - f^2/N^2$ (which is very nearly unity), their equation is of a very similar form to the equation used by Senft *et al.* (1993) relating the density variance spectrum to the horizontal wind variance spectrum, obtained using the 'approximate' dispersion and polarisation relations. Allen and Vincent (1995) noted that they used 'standard textbook formulations of the polarisation relations', but do not give details. The similarity between the spectral equations of Senft *et al.* (1993) and Allen and Vincent (1995) indicates that the same (approximate) polarisation relation may have been used in their derivation. As a check on this, the derivation of the relation between density (or equivalently temperature) and energy was performed, following the general outline in Section 4.8.2, using the full dispersion relation of $m^2\,(\omega^2 - f^2) = k_H^2\,(N^2 - f^2)$, but with the 'approximate' polarisation equation (i.e. Equation (3) of Senft *et al.*, 1993, given above) of $W'^2 = (k/m)^2\, U'^2$. (This derivation is reproduced in Appendix A.4.) The resulting relationship was found to be identical with that presented by Allen and Vincent (1995), implying that the 'approximate' polarisation relationship was used by them.

The errors introduced by the use of the approximate relationship are relatively minor, as they tend to average out when integrating over all frequencies to obtain the vertical wavenumber spectra, but can be significant if more limited bandwidths are considered. (See Table 3 for an indication of this.) However, as very few observational data sets are likely to be of sufficiently long duration to be able to resolve the very low frequency wave perturbations, the integration of the 'approximate' formalism over all frequencies performed by Allen and Vincent (1995) to derive their scaling factor (of approximately $^4/_3$) appears inappropriate. A scaling factor close to unity would appear a better choice (e.g. Table 2). Hence it seems possible that the power spectra calculated by Allen and Vincent (1995) may over estimate the true power by some 30%. Given the current observational uncertainties, and complications such as background winds Doppler shifting true wave frequencies to 'observed' frequencies, such an overestimate is probably not of significance, although again it is important to be aware of this point.

Based on the analysis presented above, scaling the relative temperature or density spectra by $^1/_2(g/N)^2$ (or $^1/_4(g/N)^2$) to produce total (or time averaged) energy per unit mass spectra appears the most appropriate interpretation of Rayleigh lidar data in the study of gravity wave activity. Identifying the $(g/N)^2$ scaled density (or temperature) variance with available potential energy per unit mass is mathematically valid for all frequencies, while identifying these fluctuations with kinetic energy per unit mass is valid for all but the very lowest frequencies (i.e. those comparable with the inertial frequency), and in most cases will also be acceptable. In any final presentation of lidar-derived energy per unit mass spectra it will be important to state explicitly how such spectra were obtained, and also as to whether they represent time-averaged or total energy per unit mass calculations.

4.10 Conclusions

The above derivations outline the means by which gravity wave properties can be estimated from Rayleigh lidar observations of the normalised wave-induced density (or temperature) fluctuations. It is suggested, after considering various approximations used in these derivations, that scaling the spectra of the normalised fluctuations of density or temperature by $(g/N)^2$ and interpreting the result as 'potential energy per unit mass' spectra, specifying whether these are total or time averaged quantities, is to be preferred.

4.11 Acknowledgements

The lidar project is supported by the Australian Antarctic Division and the University of Adelaide.

References

Allen, S.J. and Vincent, R.A. (1995). Gravity wave activity in the lower atmosphere: seasonal and latitudinal variations. *Journal of Geophysical Research,* 100: 1327–1350.

Collins, R.L., Tao, X. and Gardner, C.S. (1996). Gravity wave activity in the upper mesosphere over Urbana, Illinois: lidar observations and analysis of gravity wave propagation models. *Journal of Atmospheric and Terrestrial Physics,* 58: 1905–1926.

Eckermann, S.D. (1995). Effect of background winds on vertical wavenumber spectra of atmospheric gravity waves. *Journal of Geophysical Research,* 100: 14097–14112.

Fritts, D.C. (1995). Gravity wave forcing and effects in the mesosphere and lower thermosphere. In: Johnson, R.M. and Killeen, T.L. (Eds.). *The upper mesosphere and lower thermosphere: A review of experiment and theory.* Geophysical Monograph 87, American Geophysical Union, Washington. Pp. 89–100.

Fritts, D.C. and VanZandt, T.E. (1993). Spectral estimates of gravity wave energy and momentum fluxes. Part I: energy dissipation, acceleration, and constraints. *Journal of the Atmospheric Sciences,* 50: 3685–3694.

Gardner, C.S., Miller, M.S. and Liu, C.H. (1989). Rayleigh lidar observations of gravity wave activity in the upper stratosphere at Urbana, Illinois. *Journal of the Atmospheric Sciences,* 46: 1838–1854.

Gill, A.E. (1982). *Atmosphere-Ocean Dynamics.* International Geophysics Series, Vol. 30. Academic Press, San Diego.

Gossard, E.E. and Hooke, W.M. (1975). *Waves in the atmosphere.* Elsevier Scientific Publishing, Amsterdam.

Hall, C.M. and Hoppe, U.-P. (1997). Characteristic vertical wavenumbers for the polar mesosphere. *Geophysical Research Letters,* 24: 837–840.

Hauchecorne A. and Chanin, M.-L. (1980). Density and temperature profiles obtained by lidar between 35 and 70 km. *Geophysical Research Letters,* 7: 565–568.

Hines, C.O. (1960). Internal atmospheric gravity waves at ionospheric heights. *Canadian Journal of Physics,* 38: 1441–1481.

Hostetler, C.A. and Gardner, C.S. (1994). Observations of horizontal and vertical wave number spectra of gravity wave motions in the stratosphere and mesosphere over the mid-Pacific. *Journal of Geophysical Research,* 99: 1283–1302.

Klekociuk, A.R. (2001). Lidar observations of the middle atmosphere above Kingston, Tasmania (43°S, 147°E) between March 1997 and April 1999. In: Morris, R.J. and Wilkinson, P.J. (Eds.). *ANARE Reports 146.* Australian Antarctic Division. Pp. 25–38.

Klekociuk, A.R., Argall, P.S., Morris, R.J., Yates, P., Fleming, A., Vincent, R.A., Reid, I.M., Greet, P.A. and Murphy, D.J. (1994). A lidar for the study of the atmosphere above Davis, Antarctica. In: Burns, G. and Duldig, M. (Eds.). *ANARE Research Notes* 92. Australian Antarctic Division. Pp. 42–52.

Makhlouf, U., Dewan, E., Isler, J.R. and Tuan, T.F. (1990). On the importance of the purely gravitationally induced density, pressure and temperature variations in gravity waves: their application to airglow observations. *Journal of Geophysical Research,* 95: 4103–4111.

Meriwether, J.W., Dao, P.D., McNutt, R.T., Klemetti, W., Moskowitz, W. and Davidson, G. (1994). Rayleigh lidar observations of mesosphere temperature structure. *Journal of Geophysical Research,* 99: 16973–16987.

Mitchell, N.J., Thomas, L. and Marsh, A.K.P. (1991). Lidar observations of long-period gravity waves in the stratosphere. *Annales Geophysicae,* 9: 588–596.

Mitchell, N.J., McDonald, A.J., Reid, S.J. and Price, J.D. (1996). Observations of gravity waves in the upper and lower stratosphere by lidar and ozonesondes. *Annales Geophysicae,* 14: 309–214.

Salby, M.L. (1996). Fundamentals of Atmospheric Physics. *International Geophysics Series, Vol. 61.* Academic Press, San Diego. p. 394.

Senft, D.C. and Gardner, C.S. (1991). Seasonal variability of gravity wave activity and spectra in the mesopause region at Urbana. *Journal of Geophysical Research,* 96: 17229–17264.

Senft, D.C., Hostetler, C.A. and Gardner, C.S. (1993). Characteristics of gravity wave activity and spectra in the upper stratosphere and upper mesosphere at Arecibo during early April 1989. *Journal of Atmospheric and Terrestrial Physics,* 55: 425–439.

VanZandt, T.E. (1982). A universal spectrum of buoyancy waves in the atmosphere. *Geophysical Research Letters,* 9: 575–578.

Whiteway, J.A. and Carswell, A.I. (1994). Rayleigh lidar observations of thermal structure and gravity wave activity in the high arctic during a stratospheric warming. *Journal of the Atmospheric Sciences,* 51: 3122–3136.

APPENDIX A

A1 Derivation of the equation: $|X|^2/|R|^2 = (N^2 - \omega^2)\ (g^2/N^4)$ from Section 4.5.2

We start with the gravity wave polarisation equations from Hines (1960), listed as Equation (1) above, with the individual terms described in Equation (2). Forming $|X|^2$, and noting $k = \omega/v$ as before, we have

$$|X|^2 = \omega^4/v^2C^4(m^2 + (1 - \gamma)^2g^2/C^4).$$

Using the gravity wave dispersion relation (Equation (1) above), viz. where

$$m^2 = (N^2 - \omega^2)/v^2 - (\omega_a^2 - \omega^2)/C^2$$

to substitute for m, and taking out a factor of $1/v^2$, we have

$$|X|^2 = \omega^4C^4/v^4((N^2 - \omega^2) - (\omega_a^2 - \omega^2)(v^2/C^2) + (1 - \gamma)^2g^2(v^2/C^4)).$$

For the case where $(v/C)^2 \sim 0$, as discussed in Section 4.5.2, this simplifies to

$$|X|^2 = \omega^4/v^4C^4(N^2 - \omega^2) \qquad (A.1)$$

For $|R|^2$, we have,

$$|R|^2 = \omega^4m^2 + (\gamma - 1)^2g^2k^4 - \gamma(\gamma - 1)g^2k^2\omega^2/C^2 + \gamma^2g^2\omega^4/4C^4,$$

as before. As $k = \omega/v$, we have:

$$|R|^2 = \omega^4m^2 + (\gamma - 1)^2g^2\omega^4/v^4 - \gamma(\gamma - 1)g^2\omega^2\omega^2/(v^2C^2) + \gamma^2g^2\omega^4/4C^4.$$

Collecting the common factor of ω^4, and using $\omega_a = \gamma g/2C$, and $N^2 = (\gamma - 1)g^2/C^2$ gives:

$$|R|^2 = \omega^4\{m^2 + (\gamma - 1)^2g^2/v^4 - \gamma N^2/v^2 + \omega_a^2/C^2\}.$$

Substituting for m^2 using the dispersion relation gives:

$$|R|^2 = \omega^4\{(N^2 - \omega^2)/v^2 - (\omega_a^2 - \omega^2)/C^2 + (\gamma - 1)^2g^2/v^4 - \gamma N^2/v^2 + \omega_a^2/C^2\}.$$

Taking out a factor of $1/v^4$, and collecting like terms gives:

$$|R|^2 = (\omega^4/v^4)\{v^2N^2 - \omega^2v^2 + \omega^2v^4/C^2 + (\gamma - 1)C^2N^2 - \gamma N^2v^2\},$$

so, taking out a factor of C^2, and collecting terms once more, gives:

$$|R|^2 = (\omega^4/v^4)C^2\{N^2[(1 - \gamma)v^2/C^2 + (\gamma - 1)] - \omega^2v^2/C^2 + \omega^2v^4/C^4\}.$$

Again, let $(v/C)^2 \sim 0$, so we are left with:

$$|R|^2 = (\omega^4/v^4)C^2N^2(\gamma - 1). \qquad (A.2)$$

Hence, dividing (A.1) by (A.2) gives $|X|^2/|R|^2 = [C^4(N^2 - \omega^2)]/[C^2N^2(\gamma - 1)]$, or more simply:

$$|X|^2/|R|^2 = (C^2/N^2(\gamma - 1))(N^2 - \omega^2),$$

and as $C^2 = (\gamma - 1)g^2/N^2$, we have

$$|X|^2/|R|^2 = (g^2/N^4)(N^2 - \omega^2) \qquad (A.3)$$

This is the Equation (12), given in Section 4.5.2. Note that this equation does not take into account the effects of the Earth's rotation. This is however simple to include - see the last part of Appendix A.2.

A2 Derivation of Equation (28) from the equations given by Gill (1982)

Gill (1982), in Section 8.5 (pp. 262–266) considers the effect of rotation of the Earth on the polarisation relations for internal gravity waves. In symbols used here, Equation 8.5.2 of Gill is:

$$W' = -(m\omega)(\Delta p/\rho_0)/(N^2 - \omega^2).$$

This relates the vertical velocity perturbation, W', to the ratio of the change in pressure to the mean density, for a given wave induced fluctuation. Similarly, for the two horizontal velocity components, U_x' (x-direction) and U_y' (y-direction) we have, from Gill's Equation 8.5.3:

$$U_x' = (k\omega + ilf)(\Delta p/\rho_0)/(\omega^2 - f^2);$$

and

$$U_y' = (l\omega - ikf)(\Delta p/\rho_0)/(\omega^2 - f^2),$$

where k and l are the horizontal wave numbers in the x and y directions. Note that now rotational effects are being considered the horizontal motion is no longer isotropic, hence the introduction of the y-coordinate. If we consider the square of the modulus of the total horizontal velocity perturbation however, $U'^2 = U_x'^2 + U_y'^2$, we have:

$$U'^2 = (k^2\omega^2 + l^2f^2 + l^2\omega^2 + k^2f^2)(\Delta p/\rho_0)^2/(\omega^2 - f^2)^2.$$

Writing $k_H^2 = k^2 + l^2$, the total horizontal wavenumber, and substituting for $(\Delta p/\rho_0)^2$ using the expression for W' above, we obtain:

$$U'^2 = (k_H/m)^2(\omega^2 + f^2)(N^2 - \omega^2)^2/[(\omega^2 - f^2)^2\omega^2]W'^2,$$

then with use of the dispersion relation (Equation (27)), we have:

$$U'^2 = [(\omega^2 + f^2)(N^2 - \omega^2)]/[(\omega^2 - f^2)\omega^2]W'^2.$$

This is Equation (28) quoted above.

Note that it is straightforward to extend the relation between horizontal velocity and density perturbations, derived in Appendix A.1 above, to the case including the rotational effects of the Earth's atmosphere using Equation (28). From Equation (30) derived in Section 4.8 we have $W'^2 = \rho'^2\omega^2g^2/N^4$. (This is valid for the case of a rotating atmosphere, and

is also given as Equation 8.5.6 of Gill, 1982.) Substituting in for W'^2 in Equation (28) gives:

$$U'^2 = (g^2/N^4)[(\omega^2 + f^2)(N^2 - \omega^2)/(\omega^2 - f^2)]\rho'^2,$$

which was given as Equation (31). This of course simplifies to the expression for the non-rotating case, when $\omega^2 >> f^2$, which was derived above, viz.:

$$U'^2 = (g^2/N^4)(N^2 - \omega^2)\rho'^2.$$

A3 Equality of Equation (39) and Equation 8.6.6 of Gill (1982)

Equation 8.6.6 of Gill (1982) is:

$$KE/PE = (\omega^2 + f^2\sin^2\phi')/(\omega^2 - f^2\sin^2\phi').$$

Equation (39) above is:

$$<PE>/<KE> = (\omega^2 - f^2)/(\omega^2 + f^2). \qquad (39)$$

This second equation indicates that for very low frequency waves the ratio of potential to kinetic energy is very small.

Note: $<PE_V> = {}^1/_4\rho_0(g/N)^2\,\rho'^2$, or $<PE_M> = {}^1/_4(g/N)^2\rho'^2$.

From Section 4.8.2 we had

$$<KE_M> = {}^1/_4(g^2/N^2)\rho'^2(\omega^2 + f^2 - 2\omega^2f^2/N^2)/(\omega^2 - f^2). \qquad (A.4)$$

With the approximation that $f^2/N^2 \sim 0$ this equation simplifies to Equation (34):

$$<KE_M> = {}^1/_4(g^2/N^2)\rho'^2(\omega^2 + f^2)/(\omega^2 - f^2).$$

Clearly, this equation divided by the expression for $<PE_M>$ gives Equation (39), and not Equation 8.6.6 of Gill (1982). If however we do not make this approximation at this stage, we can derive the following: from the definition of ϕ' (Equation 8.4.15 of Gill, 1982):

$$\omega^2 = f^2\sin^2\phi' + N^2\cos^2\phi',$$

we substitute for ω^2 in the third term of the numerator of (A.4) to obtain:

$$<KE_M> = {}^1/_4(g^2/N^2)\rho'^2(\omega^2 + f^2 - 2(f^2\sin^2\phi' + N^2\cos^2\phi')f^2/N^2)/(\omega^2 - f^2).$$

Making the approximation that $f^4 << f^2$ we have:

$$<KE_M> = {}^1/_4(g^2/N^2)\rho'^2(\omega^2 + f^2 - 2f^2\cos^2\phi')/(\omega^2 - f^2).$$

Using the trigonometric identity of $\sin^2\phi' - \cos^2\phi' = 1 - 2\cos^2\phi'$ this becomes:

$$<KE_M> = {}^1/_4(g^2/N^2)\rho'^2(\omega^2 + f^2\sin^2\phi' - f^2\cos^2\phi')/(\omega^2 - f^2).$$

Now, Gill (1982) has, from the dispersion relation, Equations 8.4.16:

$$N^2 - \omega^2 = (N^2 - f^2)\sin^2\phi', \text{ and } \omega^2 - f^2 = (N^2 - f^2)\cos^2\phi'.$$

Using the latter expression to replace the denominator, and the expression defining ϕ' (Equation 8.4.15 of Gill) to substitute for ω^2 in the numerator, we obtain:

$$<KE_M> = {}^1/_4(g^2/N^2)\rho'^2[(N^2 - f^2)\cos^2\phi' + 2f^2\sin^2\phi')]/[(N^2 - f^2)\cos^2\phi'].$$

Making the approximation $N^2 >> f^2$ (i.e. $f^2/N^2 \sim 0$) gives:

$$<KE_M> = {}^1/_4(g^2/N^2)\rho'^2[(N^2\cos^2\phi' + 2f^2\sin^2\phi')]/(N^2\cos^2\phi').$$

As $\omega^2 - f^2 = (N^2 - f^2)\cos^2\phi'$ (quoted above, from Gill, Equation 8.4.16), rearranging this gives $N^2\cos^2\phi = \omega^2 - f^2(1 - \cos^2\phi') = \omega^2 - f^2\sin^2\phi'$. Hence substituting for $N^2\cos^2\phi$ in the numerator and denominator gives:

$$<KE_M> = {}^1/_4(g^2/N^2)\rho'^2(\omega^2 + f^2\sin^2\phi')/(\omega^2 - f^2\sin^2\phi').$$

As we have $<PE_M> = {}^1/_4(g/N)^2\rho'^2$, hence

$$<KE_M>/<PE_M> = (\omega^2 + f^2\sin^2\phi')/(\omega^2 - f^2\sin^2\phi').$$

This is Equation 8.6.6 of Gill (1982). Hence, by only making the approximation $f^2/N^2 \sim 0$, with the definition of ϕ', and the dispersion relation, we have shown:

$$(\omega^2 + f^2\sin^2\phi')/(\omega^2 - f^2\sin^2\phi') = (\omega^2 + f^2)/(\omega^2 - f^2).$$

The reason why the ratio $<KE_M>/<PE_M>$ differs from unity for very low frequency waves is covered in some detail in Gill (1982) - e.g. chapters 7 and 8 - in a discussion of the concept of available potential energy in a rotating system. At these low frequencies not all of the potential energy is able to be converted into kinetic energy, whereas all of the available potential energy can be. The difference arises as in the rotating atmosphere some potential energy is assimilated into the general equilibrium state, such as balancing horizontal pressure gradients, and is not available for conversion to kinetic energy. (See Gill, 1982, for more details.)

A4 Derivation of the spectral equation of Allen and Vincent (1995)

Starting with the 'full' dispersion relation of $m^2(\omega^2 - f^2) = k_H^2(N^2 - f^2)$, the 'approximate' polarisation equation of $W'^2 = (k/m)^2U'^2$, and the expression for kinetic energy per unit mass of $<KE_M> = {}^1/_4(U'^2 + W'^2)$, we have (with $k_H = k$):

$$<KE_M> = {}^1/_4[(N^2 - \omega^2)/(\omega^2 - f^2) + 1]W'^2,$$

$$= {}^1/_4[(N^2 - f^2)/(\omega^2 - f^2)]W'^2.$$

As $W'^2 = \rho'^2\omega^2g^2/N^4$, and total energy $E_T = 4<KE_M>$ (in the absence of rotation), we have:

$$E_T = [(N^2 - f^2)/(\omega^2 - f^2)]\rho'^2\omega^2g^2/N^4.$$

Working with the term in square brackets, and taking a common factor of N^2 from the numerator, and a factor of ω^2 from the denominator, and cancelling, we have:

$$E_T = [1 - (f/N)^2]/[1 - (f/\omega)^2](g/N)^2\rho'^2,$$

or on transposing, and knowing that $\rho'^2 = T'^2$:

$$T'^2 = (N/g)^2[1 - (f/\omega)^2]/[1 - (f/N)^2]E_T.$$

This is the form of the spectral equation used by Allen and Vincent (1995).

5. SIMULATIONS OF LIDAR OBSERVATIONS OF VERTICAL WAVENUMBER SPECTRA OF ATMOSPHERIC GRAVITY WAVES

J.L. Innis, A.R. Klekociuk and D.J. Murphy

Atmospheric and Space Physics
Australian Antarctic Division
Kingston Tasmania 7050 Australia
(email: john.innis@aad.gov.au)

Abstract

Middle-atmosphere gravity wave spectra of energy against vertical wavenumber are usually assumed to have a log-log slope of -3, based on a number of studies over many years and geographical locations, as well as from theoretical work. A series of simulation runs was performed in order to investigate the criteria required for the satisfactory recovery of the input gravity wave spectrum from lidar data, in the face of the typical signal-to-noise ratios and also from effects due to the waves themselves. The simulations involved taking a model atmospheric density profile, converting to equivalent lidar photon counts by scaling by the inverse square of height and adding photon shot noise corresponding to the photon counts for a given height. To this basic 'lidar profile' a number of waves with characteristics typical of middle atmosphere gravity waves were added: the amplitude of the waves were set to be proportional to $m^{-1.5}$ or m^{-1}, hence yielding wave populations with energy per unit mass dependencies (proportional to the amplitude squared) of either m^{-3} or m^{-2}. The wave frequencies were chosen at random from the interval $\omega = f$ to $\omega = N$, but from a population with an $\omega^{-5/3}$ distribution.

Sets of simulations were performed with either 20 or 200 such waves added to the basic lidar profile, for three different values of signal-to-noise ratios. These synthetic 'added wave' density profiles were then treated in the same way as real lidar data is analysed, in order to produce profiles of relative density fluctuations as input into harmonic analysis routines. The resulting slopes measured from the calculated power spectra showed, in general, that the analysis procedures used were able to recover the input spectral slope, but often with an error of order of ±0.5 or ±1.0 (i.e. -3.0 ±0.5, or -3.0 ±1.0), for an individual simulated data set of effective length comparable to a typical night's data. The mean of ten such simulations recovered the input slope reasonably well for the cases when 200 monochromatic gravity waves were added to the data set, and less well for the case of 20 added waves. The results suggest that a gravity wave spectral slope measured from an individual night's observation should be interpreted cautiously, as a number of nights of data are required in order to characterise both the mean slope and the observational scatter in the measurement.

5.1 Introduction

Vertical wavenumber spectra of energy per unit mass (or related quantities) of atmospheric gravity waves have been determined from a variety of observational methods for a number of years. The main features resulting from these studies is a near 'universal' spectrum, largely independent of season and location, with a spectral slope (on a log-log plot) near -3 for vertical wavenumbers less than the 'characteristic' vertical wavenumber, $m = m_*$, where the slope turns over to become near zero or +1. The theoretical description of gravity wave spectra has also received much attention (e.g. VanZandt, 1982; Fritts and VanZandt, 1993; Zhu, 1994). The characteristic vertical wavenumber shows an altitude dependence, as m_* decreases with height. At mesosphere heights, estimates of m_* range from around 4×10^{-4} rad m^{-1}, equivalent to a characteristic vertical wavelength of 16 km (Hall and Hoppe, 1997), to around 2×10^{-4} rad m^{-1} (Fritts, 1995). Several recent reviews of middle atmosphere gravity waves are available (e.g. Fritts, 1995; Hocking, 1996).

Lidar observations are also being used to obtain gravity wave spectra. A number of groups around the world are active in this field (e.g. Collins *et al.*, 1996; Meriwether *et al.*, 1994; Mitchell *et al.*, 1996; Whiteway and Carswell, 1994). Usually, lidar measurements of atmospheric density or inferred temperature fluctuations are used, via the gravity wave polarisation equations (e.g. Hines, 1960; Gill, 1982) to transform these data into gravity wave energy per unit mass fluctuations, and hence to derive gravity wave spectra. A companion paper (Innis *et al.*, 2001) gives the background to these methods.

In this paper, we perform a series of spectral calculations of some simulated gravity wave data, and explore some of the criteria required of the data set to lead to a satisfactory recovery of the input spectrum. We also outline how the gravity wave spectra are calculated from the raw lidar data. These calculations are intended as an aid in interpreting gravity wave spectra to be obtained from the operation of the Australian Antarctic Division lidar (Klekociuk *et al.*, 1994). This lidar is currently undergoing test observations at the Antarctic Division headquarters in Kingston, Tasmania, prior to deployment at Davis station in early 2001.

5.2 Calculation of vertical wavenumber spectra

As discussed in Innis *et al.* (2001), vertical wavenumber spectra of the time averaged potential energy per unit mass of the gravity waves can be found from:

$$E_{PEM} = {}^1/_4 (g/N)^2 E_\rho,$$

where E_{PEM} is the spectrum required, g is the acceleration due to gravity, N is the Brunt-Vaisala frequency, and E_ρ is the vertical wavenumber spectrum of the normalised density fluctuations due to the gravity waves. That is, the density variations are found from:

$$\Delta\rho = \rho_i - \rho_o$$

where ρ_i is the i^{th} density profile, and ρ_o is the mean density profile for the night. Both ρ_i and ρ_o are of course functions of altitude, i.e. $\rho_i(z)$ and $\rho_o(z)$. The height dependence for these and other quantities is implicit in all of the following. The variations are normalised by dividing by the mean density, to form the normalised density perturbations, ρ':

$$\rho' = \Delta\rho/\rho_o.$$

The power spectrum of ρ' is found, via an FFT algorithm. This spectrum is then scaled by multiplying by $^1/_4(g/N)^2$ to obtain the vertical wavenumber spectrum of time averaged potential energy per unit mass. The *m*-spectrum of energy per unit mass of gravity waves usually exhibits a -3 power law, referred to above. (See Innis *et al.*, 2001 and references therein for further details.)

Hence each density profile yields an estimate of the gravity wave potential energy per unit mass spectrum. As a number of density profiles are obtained during a typical observing run, the individual spectra obtained are averaged in order to reduce the errors involved in any single measurement.

Both g, the acceleration due to gravity, and N, the Brunt-Vaisala frequency, are functions of altitude, and must be calculated for each height used in the ρ' calculation. The calculation of g is straightforward, from Newton's law of universal gravitation:

$$g(z) = GM_e/(r_e + z)^2,$$

where G is the universal gravitational constant, M_e is the mass of the Earth, r_e is the radius of the Earth, and z is the height under consideration. No allowance is required for the non-sphericity of the Earth, or other higher order effects, as errors introduced by ignoring these terms are much less than the observational uncertainties in density and temperature.

Calculation of N proceeds in the following manner. Firstly the mean temperature profile with height for the data set is calculated, from the density measurements, using the standard lidar technique (e.g. Hauchecorne and Chanin, 1980). This assumes hydrostatic equilibrium and the ideal gas law. The Brunt-Vaisala frequency is found from (Hargreaves, 1992):

$$N(z) = g(z)[M(\gamma - 1)/(\gamma RT(z))]^{0.5},$$

where g(z) is the gravitation acceleration as before, M is the air mean molecular mass (approximately 2.895×10^{-3}, in SI units - kg mol^{-1}, in the middle atmosphere), γ is the ratio of specific heats (approximately 1.4 up to 90 km altitude), R is the gas constant, and T(z) is the temperature profile derived from the density observations. This equation is derived from the more familiar $N^2 = (\gamma - 1)g^2/C^2$ (e.g. Hines, 1960).

An alternative calculation of N is possible using the equation:

$$N(z)^2 = [g(z)/T(z)][\delta T(z)/\delta z - \gamma],$$

where $\delta T(z)/\delta z$ is the temperature gradient with height in the atmosphere. However, as T(z) often has errors of order of several degrees K for a measurement at a given height, the measurement of the gradient $\delta T(z)/\delta z$ is subject to large point-to-point excursions. Some relatively severe smoothing of the $\delta T(z)/\delta z$ profile would be required. Hence it is felt that the earlier expression for N is a more appropriate choice in the circumstances.

Thus the lidar data supply not only the density observations used to calculate the energy spectrum, but also the background temperature profile, T(z), required for the determination of N.

5.3 The simulations

5.3.1 The model density profiles

In order to provide a skeleton synthetic lidar data set for use in the simulations, a series of synthetic lidar photon count versus altitude profiles were constructed. The simulated photon count profiles were placed into a lidar data structure, replacing the photon counts observed on the night of 12 September 1997 with the Australian Antarctic Division lidar during a test observation at Kingston, Tasmania. On this night density profiles were obtained at 30 second intervals, with a vertical resolution of 93.69 m. These data were later binned by a factor of four in time, and five in vertical resolution. A total of 86 usable profiles resulted. Hence, the synthetic lidar data set used in these simulations consisted of 86 sequential synthetic profiles, with temporal and vertical resolutions of 120 seconds and 468.44 m respectively. These spatial and temporal limits were chosen on the basis that the upper frequency limit for gravity waves, N, is around 4 minutes in the middle atmosphere, and the vertical wavenumber slope of -3 means that gravity waves with vertical wavelengths less than 1000 m are a relatively small component of the spectrum.

The simulated photon counts were produced as follows. The MSISE-90 model atmosphere (Hedin, 1991) was used to calculate a model density profile for the height range 0 to ~ 90 km. The density profile was turned into a simulated lidar photon count profile, using the relation

$$\text{photon counts } (z) = K \times \text{density}(z)/z^2.$$

where the constant K was chosen in order to produce simulated photon count rates comparable with the maximum count rates obtained by the lidar. Photon noise was added to each profile, based on a Gaussian distribution of the square root of the number of counts in any given altitude bin for any given profile. (i.e. in bin x, with count rate c, the added noise was selected from a Gaussian distribution of mean of zero and standard deviation $\sqrt{c}$. This added noise could thus be 'negative'.) Thus, for the 86 initial profiles, each differed from the others only as a result of this added noise.

This synthetic data set was used as the base data for all of the following simulations. To these synthetic profiles a series of waves were added, the characteristics of which are described in the following sections. Power spectra were calculated over the altitude range 38 to 55 km, as a representative height range that may be employed in the analysis of lidar data.

5.3.2 The wave amplitude dependence on vertical wavenumber

As stated earlier, the vertical wavenumber spectrum of potential energy per unit mass for middle-atmosphere gravity waves displays a log-log slope near -3 for vertical wavenumbers (m) greater than the characteristic vertical wavenumber $(m)_*$. In other words for a ten fold increase in vertical wavenumber, the potential energy per unit mass decreases by one thousand times in this interval.

We limit this study to waves of vertical wavenumber between $(m/2\pi) = 7.25 \times 10^{-5}$ m^{-1} to 100×10^{-5} m^{-1}. That is, the limits to vertical wavelength are from around 14 to 1 km. This wavelength range spans from near the upper limit of a wave detectable with the lidar data (around 15 km for a representative maximum 30 km data range), down to the lower limit set by the vertical resolution used in the synthetic data set. Hence we are also effectively limiting ourselves to the case where $m > m_*$. Note that as the m range spans around -1.3 in the log the relative energy change over this span, for a spectral slope of -3, is 3.9 in the log or nearly 8000 times. (In the following we quote the vertical wavenumbers in units of inverse metres, rather than the more correct radians m^{-1}. This is done in order to facilitate conversion to vertical wavelengths from a simple reciprocal, without involving the factor of 2π.)

We reproduce the -3 log spectral slope by assuming that the amplitude of a given density perturbation, due to a wave of vertical wavenumber m, is proportional to $m^{-1.5}$. As the energy of a wave is proportional to the square of the amplitude, the energy spectrum will thus be proportional to m^{-3}. Hence the low vertical wavenumber (large wavelength) waves will be significantly more energetic than the high vertical wavenumber (short wavelength) waves. Specifically, a wave of vertical wavenumber equal to 7.25×10^{-5} m^{-1} will have approximately 31.6 times the perturbation amplitude, in our simulations, of a wave of vertical wavenumber 7.25×10^{-4} m^{-1}. Also, the perturbation amplitude of this wave of vertical wavenumber 7.25×10^{-5} m^{-1} will have the maximum amplitude of all the waves used in our simulations – other waves of different m are scaled to this value. This maximum value is arbitrarily set to one of three values, as discussed in the next section.

The dependence of the size of the perturbation on vertical wavenumber was used to produce a set of waves, of vertical wavenumber randomly distributed over the limits set above, but with an amplitude set by the $m^{-1.5}$ dependency. It was these waves, with random frequencies and

phases as discussed below, that were added to the synthetic photon count profiles produced from the model densities.

An alternative interpretation of the -3 slope in the energy per unit mass spectra, in terms of relative frequency of occurrence of waves of a given vertical wavenumber, was also considered. That is, it was postulated, for the sake of argument, that all waves had a similar perturbation amplitude, regardless of vertical wavenumber, and that the greater energy at low wavenumbers indicated that more low wavenumber than high wavenumber waves were observed at any one time. Such an interpretation was found to be untenable. Unless an extremely large number of waves were assumed to be present at each instant in time, it was not possible to sample a representative range of the vertical wavenumbers known to be present in the real atmosphere. However, if such a large number of waves were present, the average fluctuation in a given perturbation quantity tended to zero, from cancellation due to the random phases of the waves. (That is, in order to produce the -3 spectral slope in this case requires 1000 times more waves at $m = m_1$ than at $m = 10m_1$, if the waves have equal perturbation amplitudes.)

For the purposes of this study, a vertical wavenumber spectrum of the energy per unit mass with a log-log slope of -2 was also constructed, in this case by having a perturbation amplitude dependence of m^{-1}. This was also used in the simulations, as a comparison with the results from the more generally accepted m^{-3} dependency of the energy per unit mass spectrum.

5.3.3 The maximum perturbation amplitude

As noted, the density perturbations considered here were assumed to have an amplitude dependence of the form $A_r(m) \propto m^{-1.5}$ or m^{-1}. Hence the spectral slope in energy per unit mass, which is related to the square of the amplitude, show a m^{-3} slope for the $m^{-1.5}$ amplitude dependence, or m^{-2} for the m^{-1} amplitude form.

For a monochromatic gravity wave the amplitude of the relative density perturbation would be expected to increase with height due to the rapid decrease of atmospheric density, from conservation of energy considerations. Observations indicate, however, the gravity wave spectrum changes little in power with altitude, which is normally interpreted as a saturation effect, i.e. the gravity wave amplitudes observed are the maximum possible at any given height. Various theoretical ideas account for how the energy from the 'excess amplitude' is redistributed in the gravity wave spectrum. The point to be noted here is that, in effect, there is little change observed in the gravity wave energy with height compared with what may be expected from the decrease in density. (See, e.g. the review by Fritts, 1995.)

Relative density (or temperature) fluctuations induced by gravity waves in the middle atmosphere are generally thought to be of the order of a few percent at maximum (e.g. Whiteway and Carswell, 1994; Gardner *et al.*, 1995; Mitchell *et al.*, 1996, among many others). Hence

we have adopted three trial amplitudes, corresponding to maximum relative density perturbations of 1.0, 2.5 and 5.0% - that is the maximum amplitude of the relative density fluctuation, and is due to the wave with the lowest allowed vertical wavenumber of $m_{min} = 7.25 \times 10^{-5}$ m^{-1}. Waves of greater m have a lesser amplitude, corresponding to the m-amplitude dependency. Hence, with the $m^{-1.5}$ dependence, a wave of $m = 4m_{min}$ has an amplitude of only $(4)^{-1.5} = (1/8)$ of the amplitude of an $m = m_{min}$ wave. For the case where the maximum amplitude is 5.0% (for $m = m_{min}$), an $m = 4m_{min}$ wave has a relative density fluctuation amplitude of $5/8$%.

Each maximum amplitude is trialled individually. The amplitude does not increase with height, but stays at this constant value of a specified percentage of the relative density fluctuation. (Note that density itself decreases rapidly with height, so that the absolute density change for a given wave perturbation also decreases rapidly.)

When adding the effects of a wave of the specified amplitude of relative density to a given synthetic photon count profile, no other noise is added in, as each profile already exhibits photon noise statistics appropriate to photon signal. As the waves are only a few percent of the relative density, at the height of interest the wave perturbation is often of comparable size to the photon noise. Hence it was thought that no extra validity would be gained by adding additional noise to the profile.

5.3.4 The number of waves added to the model profiles

Three different values of the number of waves added to each profile were trialled: 2, 20, and 200. The waves appear in each profile, shifted in phase due to the time difference from one profile to the next. (See section 5.3.5). In the simulations, the case of 2 added waves is considered only for one selected amplitude value, that of 5.0%, as the resulting spectra are very poorly characterised by a slope derived from a linear fit, which would also be the case for the other values of the maximum amplitude.

5.3.5 Temporal coherence of the added waves

Each profile of the data set, nominally of an integration of 120 seconds, was treated as one observation in a nightly sequence. As the data set consists of around 90 profiles, the total time duration represented is around 10 000 seconds.

In a similar manner to the selection of vertical wavenumber for each wave, a wave frequency was also selected at random, from a population ranging from 20×10^{-5} to 400×10^{-5} Hz (corresponding to periods from 5000 to 250 seconds), but with a relative frequency of occurrence giving a $-5/3$ slope in the log-log domain, to represent the $\omega^{-5/3}$ dependence of the frequency spectra. The frequency limits were chosen to range from the lowest period that could be detected in the data set (i.e. half of the data train length) to twice the Brunt-Vaisala period ($1/2N$ is equivalent to ~ 200 seconds in the middle atmosphere). The wave frequencies were

used to determine the time development of the individual waves, from profile-to-profile, throughout the total time represented.

In addition, for each wave, an initial phase term, ϕ, was chosen from a random uniform distribution over the interval zero to 2π.

Hence a given wave can be represented by:

$$w(z) = A_\rho m^{-Q} \sin [2\pi z/\lambda_z - \omega t + \phi],$$

where z is the height, A_ρ is the maximum amplitude, of either 1.0, 2.5 or 5.0% depending on the simulation run, m is the vertical wavenumber, λ_z is the vertical wavelength (= $1/m$ where m is in units of m^{-1}), and ω is the angular frequency of the wave, selected as discussed above. The index Q is 1.5 or 1, as discussed earlier.

Note that each profile, which contains either 2, 20 or 200 added waves, differs from the others only due to the time evolution of the waves – the vertical wavenumbers, wave amplitudes, and wave frequencies remain the same from profile-to-profile. It was hoped that this would reflect, in part at least, the likely temporal coherence of gravity waves in the atmosphere.

5.3.6 Measurement of spectral slope

The spectral slopes of the final spectra are measured by the simple device of fitting a straight line to the power spectrum between $\log(m) = -3.8$ and -3.2 (i.e. between 16×10^{-5} and 63×10^{-5} m^{-1}, or from ~ 6.3 to ~ 1.6 km). These limits were chosen to be well away from the upper limit in vertical wavelength of the input waves used in this simulation (20 km, or 5×10^{-5} m^{-1} in wavenumber), and to avoid the high wavenumber region of the spectrum where photon noise dominates, as waves of high wavenumber make a significantly smaller contribution to the density fluctuations. In the spectral plots presented here the fitted line is shown. Generally, the fit appears to give a reasonable estimate of the spectral slope in this vertical wavenumber region.

5.4 Results

For the two trial spectral slopes (-2, -3), simulations were performed using wave amplitudes of 1.0, 2.5, and 5.0%, and with 20, and 200 added waves. For the largest amplitude of 5.0% an additional set of simulations was performed for the case of two added waves. For each set of input conditions, 10 individual trials were conducted, where random waves were added to the initial 'pure' photon count profiles. Hence a total of 140 individual trials were conducted.

These 'added wave' profiles were then processed in the same way as Australian Antarctic Division lidar data are processed. A fitted background is subtracted from all profiles. The photon counts were scaled to densities via comparison with the MSISE-90 model in a 'clear signal region' (above the aerosol layers, and below the height where falling count rates limit the precision). A mean nightly temperature profile was

determined from the measured densities (to be used in the calculation of the Brunt-Vaisala frequency). Relative density fluctuations, ρ', were formed, as discussed in Section 5.2, and the ρ' spectra found. These are scaled to (time averaged) potential energy per unit mass spectra as explained above. The (log-log) slope of the resulting spectra were determined – these form the data set which will be discussed shortly.

As a check as to how well this procedure recovers the input waves, Figure 1 shows a comparison between the input waves from one such trial and the 'recovered waves', after fitting to the model density, etc. The agreement is relatively good. Some small differences are apparent – some high frequency information appears lost for example. However the general form of the recovered waves matches the input data quite well.

Also as a check on the noise background, Figure 2 shows the power spectrum of the raw (photon noise added) model density profiles without any added waves. The spectral slope was 0.08, not significantly different from zero. This spectrum can be taken as an indication of the noise floor of all spectra calculated here.

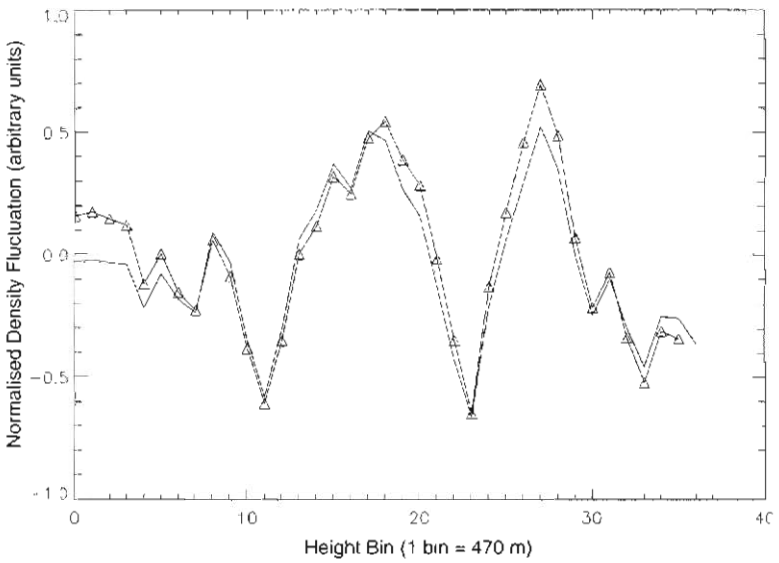


Figure 1. *A comparison between the input profile (formed by the sum of 200 waves) – shown as the solid line – which was then combined with the synthetic atmospheric density profile to form a simulated data set and the recovered 'residual wave-induced fluctutations' – solid line with triangular points – after processing with the lidar data analysis software to remove the mean density profile. The vertical axis is normalised density fluctuation, while the horizontal axis is in height bins (1 bin is about 470 m). The (simulated) height range shown corresponds to ~ 38 to ~ 55 km. The good correspondence between the two lines shows that the input waves were well recovered by the routines used.*

As a further check on the validity of the methods used, and as a preliminary to the simulations, a series of 'wave only' profiles were produced, in the manner described above, but were not combined with model density profiles. For these wave-only profiles ten trials of 200 added waves (drawn from the same wave population as used in the simulation runs reported below i.e. for input log-log slopes of both -2 and -3) were separately processed, and the resulting gradients of the *m*-spectra of energy per unit mass determined. The results of this test are shown in Table 1, where the mean and standard error in the mean of the recovered gradients are presented. The original spectral slopes appear satisfactorily recovered.

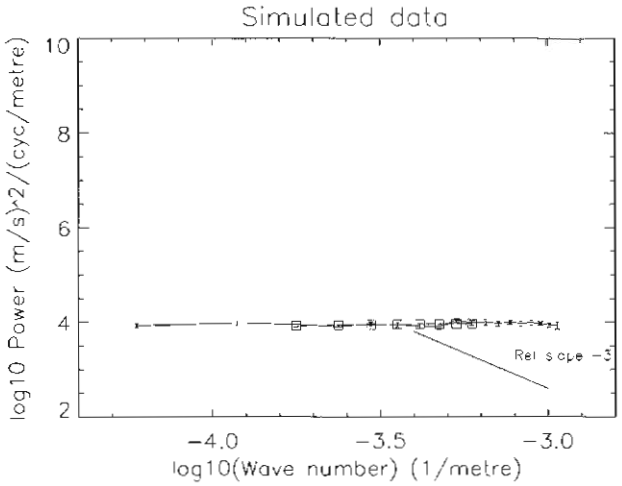


Figure 2. *Power spectrum of a series of raw (with added photon noise) model density profiles with no added waves. The recovered spectral slope is zero. The mean power level can be taken as an estimate of the photon noise floor in the simulations.*

Table 1. *Recovered gradients of the m-spectra of energy per unit mass for 200 added waves only (without an underlying density profile), drawn from a population with an energy per unit mass spectral slope as shown. Ten trials were performed for each input spectral slope. Mean and standard errors of the gradient of the recovered spectra are shown.*

Input spectral slope (log-log):	-2.0	-3.0
Recovered spectral slope (log-log):	-2.05 ±0.10	-2.98 ±0.11

The results of the simulations will now be presented. For the case of the 5.0% maximum amplitude variation, the slopes of the *m*-spectra of energy per unit mass will be presented for each of the 10 trials for the cases of 2, 20 and 200 added waves, and for input spectral slopes of -2 and -3. For the other cases, only the mean slope found from each set of 10 trials, with standard errors and maxima and mimina, will be given.

5.4.1 For the 5.0% maximum amplitude variation

The measured gradients of *m*-spectra of energy per unit mass for each of the ten trials for the given case of 2, 20 or 200 added waves, for an input slope of either -2 or -3, are given in Tables 2 and 3.

Table 2. *Results of recovered gradients of m-spectra of energy per unit mass, from an input slope = -2, for a maximum amplitude of variation of 5.0%, with the number of added waves as shown.*

Trial	200 waves	20 waves	2 waves
1	-2.50	-1.21	+1.42
2	-1.88	-0.89	-2.58
3	-2.60	-2.03	+3.60
4	-1.51	-0.64	-3.73
5	-1.07	-2.10	+0.78
6	-2.35	-2.89	-1.70
7	-1.92	-1.42	-3.30
8	-2.02	-1.31	-4.18
9	-2.07	+0.94	-1.72
10	-1.97	-1.63	+1.95
mean ±se	-1.99 ±0.14	-1.32 ±0.32	-0.95 ±0.85

From Tables 2 and 3, it is clear that for the cases where only two added waves are considered the recovered slope of the vertical wavenumber spectra of energy per unit mass shows great variation, and a mean that does not represent the input slope. This is not unexpected. The spectra obtained for these cases typically showed strong peaks, corresponding to the vertical wavenumber of one or both of the added waves. Fitting a straight line to the spectra in this case to determine the slope has little validity. Figure 3 shows a typical spectrum resulting from the case of two added waves.

For the case of 200 added waves, the mean vertical wavenumber gradient recovered is in good agreement for the case of the input slope = -2 (Table 2), and in reasonable agreement for the case of the input slope = -3 (Table 3). The agreement is somewhat worse when only 20 added waves are considered, which again is not surprising. Further discussion of these findings is given below.

A typical recovered vertical wavenumber spectrum of energy per unit mass for the case of 200 added waves, for an input spectral slope of -3, with a maximum amplitude variation of 5.0%, is presented in Figure 4. For this specific case, the measured spectral slope was -2.82.

Table 3. *As for Table 2 – results of recovered m-spectra slopes, but here for an input slope = -3, again with a maximum amplitude of variation of 5.0%, with the number of added waves as shown.*

Trial	200 waves	20 waves	2 waves
1	-2.91	-2.79	-3.30
2	-2.77	-2.60	-3.24
3	-2.00	-3.54	-0.79
4	-2.95	-2.83	-2.37
5	-2.28	-0.81	-0.12
6	-3.58	-2.64	-0.09
7	-2.18	-3.37	+3.87
8	-2.65	-1.28	+4.64
9	-2.89	-2.27	+3.80
10	-3.06	-2.24	+1.78
mean ±se	-2.73 ±0.15	-2.44 ±0.27	+0.42 ±0.94

Table 4. *Results from trials with a maximum amplitude of variation of 2.5% measured gradients of vertical wavenumber spectra of energy per unit mass.*

	Input Slope = -2		Input Slope = -3	
Quantity	200 waves	20 waves	200 waves	20 waves
|minimum slope|	-0.94	-0.35	-2.12	-1.21
|maximum slope|	-2.28	-3.10	-3.16	-3.71
mean slope	-1.85 ±0.14	-1.60 ±0.33	-2.70 ±0.11	-2.66 ±0.22

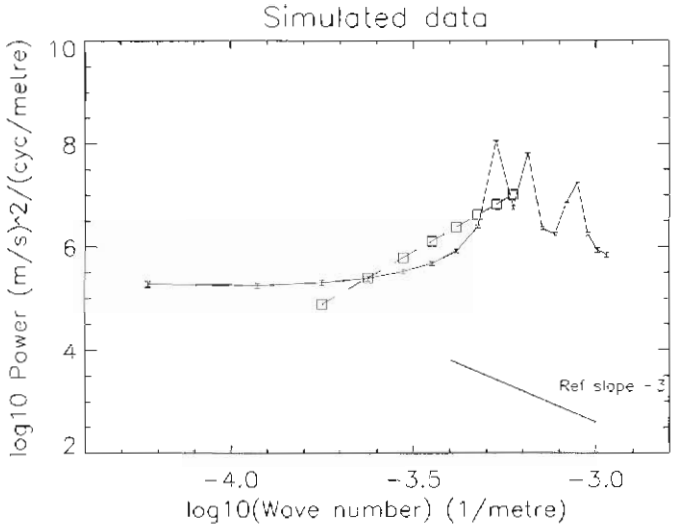


Figure 3. *A typical spectrum resulting from the case where only two waves were added into the simulated night of data. The two strongest peaks in the power spectrum, correspond to the frequencies of the added waves. The third weaker peak (about $^1/_7$ th the strength of the biggest peak) appears to be a harmonic of the first. Clearly the spectral slope determined by a linear fit (dashed line with square points) is a poor description of the spectrum in this case.*

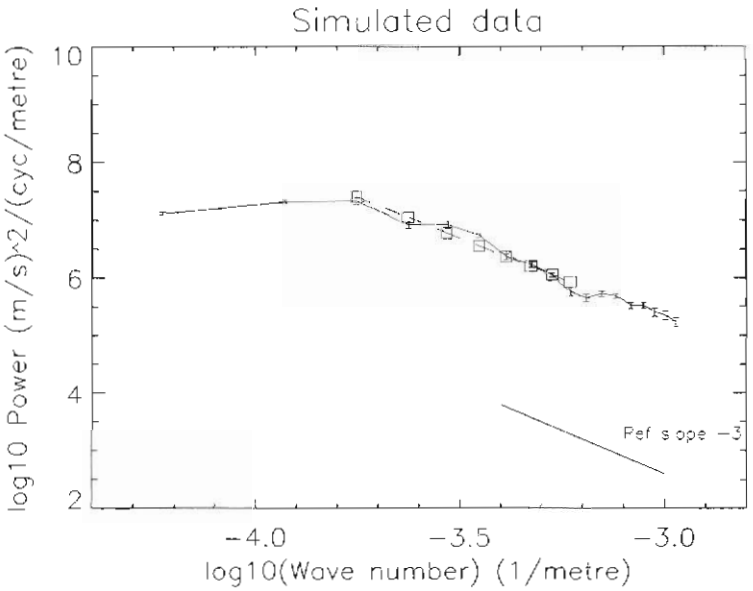


Figure 4. *A typical spectrum resulting from the trial with 200 added waves, of maximum amplitude of 5% fractional density, drawn from the population with a spectral slope of -3. Solid line - recovered spectrum. Dashed line with square points - fitted line representing the spectral slope. The recovered spectral slope in this case was -2.82.*

5.4.2 For the 2.5% and 1.0% maximum amplitude variation

The results for the trials for the case where the maximum amplitude of variation was 2.5% are presented in Table 4. These results are similar to the case where the maximum amplitude of variation was 5.0%, as presented in the previous section.

The results for the trials for the case where the maximum amplitude of variation was 1.0% are presented in Table 5. Again, the findings are similar to the previous instances.

Table 5. *Results from trials with a maximum amplitude of variation of 1.0% measured gradients of vertical wavenumber spectra of energy per unit mass.*

Quantity	Input Slope = -2		Input Slope = -3	
	200 waves	20 waves	200 waves	20 waves
|minimum slope|	-1.67	+0.15	-2.05	-1.55
|maximum slope|	-2.34	-3.04	-3.13	-3.18
mean slope	-2.05 ±0.08	-1.96 ±0.33	-2.71 ±0.09	-2.49 ±0.16

5.5 Discussion

5.5.1 General comments on the findings from these simulations

There is little effect on the recovered gradients with changing maximum amplitude of variation. This is probably a result of the fact that a five fold increase in amplitude gives only a $\sqrt{5}$ increase in signal-to-noise ratio. If anything however, results obtained here seem to show that the standard errors in the mean gradients appear to decrease with decreasing amplitude. It is possible that this is just a statistical fluctuation, as it can hardly be expected that a decreased amplitude signal could result in a more accurate measurement.

Note that with the altitude range chosen for the spectral calculations (38 to 55 km), and with the altitudinal resolution of 468 m, the number of vertical wavenumber bins in the final spectra is 18. Hence, as the input wavenumbers were selected at random, in the case of 20 added waves we would expect, on average, that only one wave would fall in any wavenumber bin. For the case of 200 waves, on average about 11 waves would fall into each bin, reducing the error in the estimate of the power in a given bin by a factor near $\sqrt{10}$.

This expectation appears borne out from the results obtained here. The standard error in the mean recovered slope decreases as the number of added waves changes from 20 to 200. The decrease is, on average, nearly a factor of 2.5. This appears consistent with the 10 times increase in the number of waves - the improvement would be expected to be

approximately $\sqrt{10}$ (~ 3.2). The deviation of the recovered spectral slope from the input slope (i.e. the 'observed' slope minus the expected slope) also appears to have a similar dependency – for the cases of 20 added waves the (observed minus expected) values are approximately twice as large as for the 200 added waves cases.

The slope = -2 case appears to be more accurately recovered than for the spectral slope = -3 case. This also appears understandable. The maximum amplitude was set to be the same for any given conditions – i.e. at either 1.0, 2.5, or 5.0% of the normalised density fluctuation – for a wave of vertical wavenumber $m_{min} = 7.25 \times 10^{-5}$ m^{-1}. In both cases waves of greater m had lesser amplitudes, falling as m^{-1} or $m^{-1.5}$, as noted above. For the latter case the energy in the waves falls at a faster rate with increasing m, hence decreasing the signal-to-noise ratio at high m more quickly than for m^{-1}. This may have tended to flatten the spectrum at high m, resulting in values of the recovered slope being less than the expected -3.

The ranges of recovered slopes in the Tables 2 to 5, even for the case of 200 added waves, show a range of around ±0.5 to ±1.0 from the expected value. Given that the synthetic data used in these simulations probably represent a 'best possible' case, as they are not prone to the various effects that are often present in real data (e.g. non-linearities, transitory dc shifts, non-photon statistic noise sources, etc.), it should not be surprising if any experiment designed to measure the slope of the gravity wave energy per unit mass spectrum in the atmosphere returns values with ranges at least of this size. Added to this are the physical processes in the atmosphere that appear to have a discernible influence on the shape of the gravity wave vertical wavenumber spectra (Eckermann, 1995). Clearly, derivation of meaningful values of spectral gradient require careful attention to the entire process of data collection and analysis, as well as an understanding of the limitations placed on the validity of the derived spectra specific to a particular instrument or data set. It also appears that a single estimate of spectral gradient, such as from one night of observation, may be of only limited value if there is no *a priori* or otherwise determinable information on the statistics of the population from which this sample measurement is taken. Long data runs will be advantageous, as always.

5.6 Conclusions

A series of simulations was performed to test the sensitivity to certain experimental parameters of the measured values of the m-spectral slope for atmospheric gravity waves. The input waves were chosen to represent the main observed features of real atmospheric waves, sampled at signal-to-noise ratios that are expected to be typical of the performance of the Australian Antarctic Division's lidar. The results reported here indicate that for waves that produce a maximum density fluctuation of a few percent of the background density, the recovered slope agrees reasonably well with the underlying spectral slope of the wave population if enough

simulated waves are included in the sample. For the cases when only a few waves are sampled, the recovered slopes are clearly only a poor approximation to the underlying slope. As it is not generally known how many waves may be sampled on any given observing night, this work indicates that spectral slopes measured from a small number of nights of lidar (or similar) data will need to be intepreted cautiously.

References

Collins, R.L., Tao, X. and Gardner, C.S. (1996). Gravity wave activity in the upper mesosphere over Urbana, Illinois: lidar observations and analysis of gravity wave propagation models. *Journal of Atmospheric and Terrestrial Physics,* 58: 1905–1926.

Eckermann, S.D. (1995). Effect of background winds on vertical wavenumber spectra of atmospheric gravity waves. *Journal of Geophysical Research,* 100: 14097–14112.

Fritts, D.C. (1995). Gravity wave forcing and effects in the mesosphere and lower thermosphere. In: Johnson, R.M. and Killeen, T.L. (Eds.). *The upper mesosphere and lower thermosphere: A review of experiment and theory.* Geophysical Monograph 87. American Geophysical Union, Washington. Pp. 89–100.

Fritts, D.C. and VanZandt, T.E. (1993). Spectral estimates of gravity wave energy and momentum fluxes. Part I: energy dissipation, acceleration and constraints. *Journal of the Atmospheric Sciences,* 50: 3685–3694.

Gardner, C.S., Tao, X. and Papen, G.C. (1995). Simultaneous lidar observations of vertical wind, temperature and density profiles in the upper mesosphere: evidence for nonseparability of atmospheric perturbation spectra. *Geophysical Research Letters,* 22: 2877–2880.

Gill, A.E. (1982). Atmosphere-Ocean Dynamics. *International Geophysics Series, Vol. 30.* Academic Press, San Diego.

Hall, C.M. and Hoppe, U.-P. (1997). Characteristic vertical wavenumbers for the polar mesosphere. *Geophysical Research Letters,* 24: 837–840.

Hargreaves, J.K. (1992). The solar-terrestrial environment. *Cambridge Atmospheric and Space Sciences series, Vol. 5.* Cambridge University Press, Cambridge.

Hauchecorne, A. and Chanin, M.-L. (1980). Density and temperature profiles obtained by lidar between 35 and 70 km. *Geophysical Research Letters,* 7: 565–568.

Hedin, A.E. (1991). Extension of the MSIS thermosphere model into the middle and lower atmosphere. *Journal of Geophysical Research,* 96: 1159–1172.

Hines, C.O. (1960). Internal atmospheric gravity waves at ionospheric heights. *Canadian Journal of Physics,* 38: 1441–1481.

Hocking, W. (1996). Small-scale dynamics of the upper atmosphere: experimental studies of gravity waves and turbulence. In: Dieminger, W., Hartmann, G.K. and Leitinger, R. (Eds.). *The Upper Atmosphere.* Springer, Berlin. Pp. 51–96,

Innis, J.L., Klekociuk, A.R. and Murphy, D.J. (2001). Derivation of gravity-wave spectra from Rayleigh lidar observations: a review of the formalism and a comparison with some published analyses. In: Morris, R.J. and Wilkinson, P.J. (Eds.). *ANARE Report 146.* Australian Antarctic Division. Pp. 39–71.

Klekociuk, A.R., Argall, P.S., Morris, R.J., Yates, P., Fleming, A., Vincent, R.A., Reid, I.M., Greet, P.A. and Murphy, D.J. (1994). a lidar for the study of the atmosphere above Davis, Antarctica. In: Burns, G. and Duldig, M. (Eds.). *ANARE Research Notes 92.* Australian Antarctic Division. Pp. 42–52.

Meriwether, J.W., Dao, P.D., McNutt, R.T., Klemetti, W., Moskowitz, W. and Davidson, G. (1994). Rayleigh lidar observations of mesosphere temperature structure. *Journal of Geophysical Research,* 99: 16973–16987.

Mitchell, N.J., McDonald, A.J., Reid, S.J. and Price, J.D. (1996). Observations of gravity waves in the upper and lower stratosphere by lidar and ozonesondes. *Annales Geophysicae,* 14: 309–214.

VanZandt, T.E. (1982). A universal spectrum of buoyancy waves in the atmosphere. *Geophysical Research Letters,* 9: 575–578.

Whiteway, J.A. and Carswell, A.I. (1994). Rayleigh lidar observations of thermal structure and gravity wave activity in the high arctic during a stratospheric warming. *Journal of the Atmospheric Sciences,* 51: 3122–3136.

Zhu, X. (1994). A new theory of the saturated gravity wave spectrum for the middle atmosphere. *Journal of the Atmospheric Sciences,* 51: 3615–3626.

6. ANALYSIS OF FABRY-PEROT SPECTRA OF LIDAR BACKSCATTER ECHOES

M. Conde

Geophysical Institute
University of Alaska
Fairbanks Alaska 99775-0800
USA
(email: mark.conde@gi.alaska.edu)

Abstract

A numerical procedure is presented for analysis of backscatter spectra recorded by an atmospheric Doppler lidar experiment. The recorded spectra are modelled by an analytic function of five parameters. The procedure solves for the set of model parameters that yields the best least-squares fit between the model and the data. The model is formulated to allow for Doppler shifted and broadened scattering from both aerosols and from multiple molecular species, for a continuum background, and for the effects of instrumental broadening. All five model parameters are fitted in the signal domain. Three parameters are fitted by a direct analytic solution and two by numerical searching. If required, one or more parameters may be held fixed at externally-supplied values, while the remainder are fitted accordingly. The method's performance was examined by applying it to numerically simulated test spectra, for several example lidar configurations. Some implications of these results for the design of actual lidar experiments are discussed.

6.1 Introduction

Lidar systems operate by illuminating the atmosphere with short duration and highly directional pulses of laser photons. Measuring properties of the backscattered photons as a function of time allows measurement of atmospheric properties as a function of distance from the lidar. A high resolution spectrometer in the receiver system can measure the wavelength spectra of the backscattered photons. The atmospheric kinetic temperature and line-of-sight wind as a function of range from the lidar can then be inferred from these spectra. The spectrometer of choice for this application is a single or dual etalon Fabry-Perot spectrometer (FPS).

Analysis of a recorded spectrum proceeds by fitting an analytically-described model spectrum to it. The model is formulated as a function of several parameters. These parameters correspond to the desired geophysical quantities (wind and temperature) plus, usually, several other geophysically useful quantities that will be estimated as a by-product.

Designing an effective analysis scheme is a three-part problem. First, the functional form of the model spectrum must be specified. Second, we need a scalar expression defining the 'goodness-of-fit' between the recorded spectrum and a particular instance of the model. Third, we need a scheme for determining the set of parameters that yields the best possible fit between the model and the recorded spectra, using this goodness estimator. For a single spectrum, the set of best-fit parameters are the experimental estimators of the desired geophysical quantities.

6.2 The model spectrum

Experimentally recorded spectra are, inevitably, instrumentally broadened by the finite spectral bandwidths of both the transmitting laser and the receiving Fabry-Perot spectrometer. However, there is no need for the analysis to represent these two broadening functions independently; a single function can correctly account for the combined effects of both. Indeed, in a typical lidar experiment, this combined effect is what is actually measured when directly viewing the transmit laser. We refer to this as the 'instrument function' and denote it mathematically with the symbol I.

It is assumed that actual lidar returns from the atmosphere will arise due to both Rayleigh scattering from molecules and Mie scattering from aerosols. (Inelastic processes, such as Raman and Brillouin scattering, will be ignored.) Because aerosol particle masses are large compared to the atmosphere's mean molecular mass, the spectral width of aerosol backscatter is much smaller than the spectral width of the Rayleigh component of the return signal. The Fabry-Perot instrument function should be chosen to be comparable to the Rayleigh Doppler width. Thus, the aerosol backscatter spectral width would appear negligible to the Fabry-Perot system, for all reasonable values of atmospheric temperature. It is assumed that aerosols drift with the neutral wind, so that the peak wavelength of the aerosol return spectrum will be the same as the peak wavelength of the molecular Rayleigh return spectrum. Finally, we assume that there will be some background continuum present in the recorded signal. These various components of the received signal must be incorporated into the model spectrum.

A scanning FPS does not record continuous spectra. Rather, a set of discrete samples of the return spectrum are obtained by stepping the FPS passband in wavelength, and measuring the received intensity at each step. (Because the instrument scans its passband in wavelength over time, care must be taken to avoid possible spectral distortions due to time variation of the return spectrum. Possible distortions due to this effect are suppressed by co-adding a large number of short duration spectral scans during each integration interval.) If each spectral scan contains N discrete samples, then at the end of the scan we will have a set of values, $\{Y_n\}$ where n varies from 0 to N – 1. The set $\{Y_n\}$ is a discrete approximation to the spectrum of backscattered laser photons, modified by the spectral response of the instrument.

From the discussion above, we will choose to model $\{Y_n\}$ with a set of samples, $\{s_n\}$ of an analytic function (of wavelength) of the form

$$S = I * B \tag{1}$$

where

I = instrument function,

B = backscatter spectrum, and

* denotes the convolution operator.

Rayleigh backscatter will occur from all atomic and molecular species in the atmosphere. The mixing ratios of these species (and therefore their relative backscatter intensities) are well known throughout the height range accessible to the lidar. Relative backscatter intensities due to the molecular species can thus appear as constants within the model. The return spectrum due to each molecular species will be modelled as a Gaussian function of wavelength. As each species has a unique mass, the Doppler width of the backscatter spectrum due to each species will be different. However, all species are assumed be at the same temperature – so only one free parameter is needed in the model to describe these different widths. All scattering species, including the aerosol, can be assumed to travel with the wind at the same velocity. Thus, again, we only need one model parameter to describe the Doppler shift.

For the purpose of describing the backscatter spectrum, the aerosol could potentially be treated as just another atmospheric species, *albeit* one of very high mass. However, we treat it separately from the molecular species, for two reasons. First, the aerosol abundance is variable, yielding a relative backscatter intensity that cannot be described effectively by an *a-priori* constant. Second, the bandwidth of the aerosol return spectrum would (for a reasonably optimised experiment) be narrow compared to our wavelength sampling interval. Thus, we cannot adequately represent the aerosol return spectrum using the actual experiment's wavelength sampling interval. Rather, we must use a model that only ever requires evaluation of the convolution of the aerosol spectrum with the instrument function.

From these arguments, we can rewrite Equation (1) as the instrument function convolved with a sum of three separately treated contributions: a continuum background, aerosol Mie scattering, and Rayleigh scattering summed over all contributing molecular species. We write this as

$$S(\lambda) = \sum_{j=0}^{2} a_j S_j(\lambda) \tag{2}$$

where the a_j are the backscatter intensities of each of the three contributing terms, and

$S_0(\lambda)$ = 1 represents the continuum component,

$S_1(\lambda)$ = $I(\lambda) * \delta(\lambda - \lambda_0)$; ($\delta$ denotes the Dirac delta function)

$= I(\lambda - \lambda_0)$, the aerosol term,

$S_2(\lambda) = I(\lambda) * \sum_{k=0}^{K-1} q_k \exp - \left(\frac{\lambda - \lambda_0}{w_k(T)} \right)^2$ is the total Rayleigh contribution arising from K different molecular species,

$w_k(T)$ = is the width in wavelength units of the backscatter from the k^{th} species for a temperature of T, and

λ_0 = the Doppler shifted peak wavelength of the return signal.

We refer to the three functions $S_0(\lambda)$, $S_1(\lambda)$ and $S_2(\lambda)$ as the 'basis functions' for our model. In the expression for S_2 we have used q_k to represent the relative contribution to the Rayleigh backscatter from the k^{th} molecular species. Values for q_k are supplied to the analysis as pre-determined constants, with the property that

$$\sum_{k=0}^{K-1} q_k = 1 \tag{3}$$

Generally, the model contains five free parameters that we need to fit, i.e., a_0, a_1, a_2, λ_0 and T. However, potentially, values for some of these parameters could be supplied as constants estimated by other means. For example, above the aerosol layer, a_0, a_1 and T can be estimated directly from the height profile of the spectrally-integrated return signal. Thus, the analysis program is formulated to allow any combination of parameters to remain fixed at their externally-supplied starting values.

6.3 Defining the best fit

We use the conventional χ^2 parameter to describe the degree to which the model spectrum departs from the recorded spectrum, i.e.

$$\chi^2 = \sum_{n=0}^{N-1} \frac{(y_n - s_n)}{\sigma_n^2} \tag{4}$$

where σ_n^2 is the variance of the n^{th} element of the observed spectrum, y_n. Providing that our model accurately describes the functional form of the recorded spectrum, then the best possible choice of model parameters yields the minimum χ^2, with an expected value of N minus the number of free parameters in the model. If we denote the set of model parameters as $\{p_j\}$ then the best choice for $\{p_j\}$ occurs when

$$\frac{\partial \chi^2}{\partial p_j} = 0 \tag{5}$$

for all j. This condition provides an analytic expression that could potentially be used to calculate the optimum $\{p_j\}$. However, using the model function described by Equation (2), it is not possible to solve Equation (5) for all $\{p_j\}$.

6.4 Strategies for finding optimum values for the model parameters

In designing an analysis scheme, it is important to distinguish between the criterion used to define the optimum model parameters and the procedure used to find parameters that satisfy this criterion. So far we have described the criterion, but not the procedure.

If this is not feasible to solve Equation (5), χ^2 can be evaluated numerically for various trial (p_j) and the results used to drive a search procedure to locate the χ^2 minimum. The latter approach has usually been adopted by most Fabry-Perot spectral analysis algorithms, because it is usually not possible to solve the peak position and width parameters analytically.

The simplest search strategy, conventionally described as a 'grid search', proceeds by minimising χ^2 with regard to each model parameter in turn, while all other parameters are held fixed. Optimisation over the complete set of parameters is repeated cyclically until satisfactory convergence is obtained. This algorithm is very stable, although it can converge slowly if χ^2 variations with respect to two or more of the parameters are not independent. However, the grid search will eventually be successful even in such cases, indicating that repeated independent optimisation of each parameter is viable as a generalised strategy.

Another consequence of this is that is does not matter how χ^2 is minimised for a given parameter, nor whether the same technique is used for each one. This conclusion leads to the hybrid approach adopted for the lidar analysis problem – a grid search in which, at each iteration, some of the parameters are computed analytically and some are obtained by searching.

6.5 Analytically calculating the intensities

We can solve Equation (5) analytically for the three intensities appearing in our model, i.e. a_0, a_1, and a_2. Several algorithms exist for this. The method presented below (known as 'directly solving the normal equations') is the most straightforward to describe, and is presented to illustrate that an analytic solution does indeed exist. However, this method can be unstable if the basis $\{s_n\}$ are not clearly distinguished by the data. (By this we mean that a region exists around the global χ^2 minimum within which the values of two or more coefficients of the basis functions can be changed in a complementary fashion, so the resulting χ^2 perturbation is not statistically significant.) In the case of our model function for lidar returns, this situation could arise if the aerosol and molecular return spectra are not distinguishable within the noise on the recorded spectrum. The problem manifests itself during numerical inversion of the curvature matrix – which can become near singular. A better (but more complex) algorithm in such cases is singular value decomposition.

6.5.1 *Direct solution of the normal equations*

Substituting Equation (2) into Equation (4) gives

$$\chi^2 = \sum_{n=0}^{N-1} \frac{\left(y_n - \sum_{j=0}^{2} a_j \left(S_j \right)_n \right)^2}{\sigma_n^2} \tag{6}$$

where the notation $(S_j)_n$ refers to the j^{th} basis function evaluated at the n^{th} discrete wavelength. Equation (5) thus becomes

$$\frac{\partial \chi^2}{\partial a_i} = -2 \sum_{n=0}^{N-1} \frac{\left(S_i \right)_n \left(y_n - \sum_{j=0}^{2} a_j \left(S_j \right)_n \right)}{\sigma_n^2} = 0. \tag{7}$$

Equation (7) describes three simultaneous linear equations in three unknowns a_0, a_1, and a_2. Explicitly,

$$\sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} \left(S_0 \right)_n y_n = a_0 \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} \left(S_0 \right)_n \left(S_0 \right)_n + a_1 \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} \left(S_1 \right)_n \left(S_0 \right)_n + a_2 \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} \left(S_2 \right)_n \left(S_0 \right)_n$$

$$\sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} \left(S_1 \right)_n y_n = a_0 \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} \left(S_0 \right)_n \left(S_1 \right)_n + a_1 \sum_{n=1}^{N-1} \frac{1}{\sigma_n^2} \left(S_1 \right)_n \left(S_1 \right)_n + a_2 \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} \left(S_2 \right)_n \left(S_1 \right)_n$$

$$\sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} \left(S_2 \right)_n y_n = a_0 \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} \left(S_2 \right)_n \left(S_2 \right)_n + a_1 \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} \left(S_1 \right)_n \left(S_2 \right)_n + a_2 \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} \left(S_2 \right)_n \left(S_2 \right)_n$$

which can be expressed as

$$\sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} \left(S_i \right)_n Y_n = \sum_{j=0}^{2} \left\{ a_j \sum_{n=0}^{N-1} \left[\frac{1}{\sigma_n^2} \left(S_i \right)_n \left(S_j \right)_n \right] \right\} \tag{8}$$

for $i = 0, 1, 2$. Equation (8) can be expressed even more compactly in matrix form as

$$\mathbf{b} = \mathbf{A}\mathbf{a}, \tag{9}$$

where b is a row vector given by

$$\mathbf{b} = \left(\sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} \left(S_0 \right)_n Y_n, \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} \left(S_1 \right)_n Y_n, \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} \left(S_2 \right)_n Y_n \right) \tag{10}$$

$\mathbf{a}$ is a column vector given by

$$\mathbf{a} = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$$

and $\mathbf{A}$ is a 3 x 3 square matrix defined by

$$A_{ij} = \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} \left(S_i\right)_n \left(S_j\right)_n . \qquad (11)$$

$\mathbf{A}$ is conventionally referred to as the 'curvature matrix'.

Multiplying both sides of Equation (9) by $\mathbf{A}^I$ yields an analytic equation for the model's intensity coefficients, $\mathbf{a}$,

$$\mathbf{a} = \mathrm{b}\mathbf{A}^I. \qquad (12)$$

Although this process may appear complex, it is numerically very simple to evaluate $\mathbf{a}$. The analysis program merely constructs the row vector b and the curvature matrix $\mathbf{A}$ using the summations defined by Equations (10) and (11) respectively. A standard matrix inversion routine is then called to evaluate $\mathbf{A}^I$ the result is inserted into the very simple calculation of $\mathbf{a}$ defined by Equation (12).

There are a couple of points to note. First, these calculations must be done in the signal domain (we have not derived expressions for optimising Equation (6) analytically with respect to $\mathbf{a}$ in the transform domain). Second, values for the parameters T and λ_0 are needed to evaluate the functions S_1 and S_2. This means that initial estimates for T and λ_0 must be generated before the first attempt to solve for $\mathbf{a}$. On each subsequent iteration, the latest estimates of T and λ_0 are used. Third, as mentioned above, there may be times when one or more of the a_0, a_1 and a_2 will be determined outside of this analysis, and simply supplied as constants. Thus, a practical program will need to implement these possibilities by allowing the matrix expressions to operate for any of the cases of 3, 2 or 1 unknown(s).

6.5.2 Singular value decomposition

The singular value decomposition (SVD) algorithm has been described by Press *et al.* (1986); repetition of the description is beyond the scope of this work. SVD also provides an analytic solution for the set of coefficients $\{a_j\}$ but does so in a fashion that avoids directly inverting the curvature matrix $\mathbf{A}$ which, potentially, could be near-singular for noisy spectra. Note that SVD is not applicable for finding coefficients that appear non-linearly within our model function. Thus, SVD cannot be used to evaluate T and λ_0.

The lidar analysis was implemented using the programming language 'IDL', which includes a pre-written SVD subroutine. This implementation can return not only estimates of the coefficients $\{a_j\}$ but also estimates of their variances. (Variances of the coefficients are also available from

the normal equations, although discussion of this has been omitted above.) Note, however, that in this application there is an additional contribution to the variances of parameters $\{a_j\}$ due to the variances of T and λ_0. Analytic solutions for the linear parameters offer no way to allow for this, so it is not appropriate to use variances returned by them. Rather, we must separately compute the parameter variances simultaneously for all parameters, as is described later.

SVD is known to be significantly slower than solving the normal equations. However, tests on numerically synthesised data have shown that a PC using a 233 MHz Pentium MMX processor can run the complete analysis described here on one 128-channel spectrum in an average of around 0.45 seconds. This is considered easily fast enough, even for near-real-time data processing.

6.6 Estimating temperature and Doppler shift

6.6.1 The Levenberg-Marquardt search

Because we cannot solve Equation (5) analytically for T and λ_0, these parameters must be determined by searching. There are many possible strategies for doing this. One possibility would be to use a grid search in the Fourier transform domain, as is done by many existing routines used to analyse Fabry-Perot spectra of the airglow and aurora. This method is attractive because the number of Fourier coefficients needed to describe the data and the model function is typically a factor of ten smaller than the number of observation points in the signal domain. However, this advantage is partially lost because the grid search is very inefficient (it makes no use of the gradient of the fitting function).

The approach described below operates in the signal domain and attempts to use as much analytic information as possible. Although we cannot solve Equation (5) analytically, we can compute analytic expressions for the partial derivatives of χ^2 with respect to T and λ_0. These expressions will be used to implement the Levenberg-Marquardt (LM) algorithm to find simultaneous solutions to

$$\frac{\partial \chi^2}{\partial T} = 0 \quad \text{and} \quad \frac{\partial \chi^2}{\partial \lambda_0} = 0.$$

The LM method is particularly elegant in that it implements a smoothly-varying combination of two search strategies. When the initial guess for the model parameters is a long way from optimum, the LM search uses an inefficient but numerically stable steepest-descent search. As the parameter estimates approach their solution, the search strategy changes smoothly, placing progressively more weight upon an analytic solution of a quadratic-expansion of χ^2. The two strategies differ in the method used to generate each successive new estimate for the model parameters.

The following description of the LM procedure closely follows that presented by Press *et al.* (1986). Let us denote the set of J parameters that we wish to optimise as $\{p_j\}$ to distinguish them from the set $\{a_j\}$ that we

solved above using direct analytic methods. We can also regard the set $\{p_j\}$ as a vector, in which case we denote it simply as $\mathbf{p}$. Search procedures operate by generating a sequence of estimates for $\mathbf{p}$. The steepest descent method generates a new estimate for $\mathbf{p}$ by perturbing the current one according to

$$\mathbf{p}_{m+1} = \mathbf{p}_m - \gamma \nabla \chi^2(\mathbf{p}_m) \,, \tag{13}$$

where γ is some coefficient chosen to be small enough that a single iteration does not overshoot the χ^2 minimum. Conversely, the quadratic expansion of χ^2 method generates its sequence of estimates using a relation of the form

$$\mathbf{p}_{m+1} = \mathbf{p}_m + \mathbf{D}^{\mathrm{I}}[-\nabla \chi^2(\mathbf{p}_m)] \,, \tag{14}$$

where $\mathbf{D}$ is a J x J matrix describing the second cross-partial derivatives of χ^2 with respect to the parameters $\mathbf{p}$.

Consider our expression for χ^2, Equation (4). Using the chain rule, the gradient of χ^2 with respect to $\mathbf{p}$ can be written in component form as

$$\frac{\partial \chi^2}{\partial p_j} = -2 \sum_{n=0}^{N-1} \frac{\left(y_n - s_n\right)}{\sigma_n^2} \frac{\partial s_n}{\partial p_j} \tag{15}$$

from which we define

$$\beta_j = \sum_{n=0}^{N-1} \frac{\left(y_n - s_n\right)}{\sigma_n^2} \frac{\partial s_n}{\partial p_j} = -\frac{1}{2} \frac{\partial \chi^2}{\partial p_j} . \tag{16}$$

Taking a second partial derivative gives

$$\frac{\partial^2 \chi^2}{\partial p_j \partial p_l} = 2 \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} \left[\frac{\partial s_n}{\partial p_j} \frac{\partial s_n}{\partial p_l} - \left(Y_n - s_n\right) \frac{\partial^2 s_n}{\partial p_j \partial p_l} \right] . \tag{17}$$

Press *et al.* (1986) explain that it is conventional to drop the second-order partial derivatives from the above expression – i.e. the right-hand term of the square brackets inside the summation. This is because near the solution $(Y_n - s_n)$ should be small if our model is well chosen. Further, retaining the higher derivatives is often numerically destabilising in practice. From this we define

$$\alpha_{jl} = \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} \frac{\partial s_n}{\partial p_j} \frac{\partial s_n}{\partial p_l} \approx \frac{1}{2} \frac{\partial^2 \chi^2}{\partial p_j \partial p_l} . \tag{18}$$

Now both of Equations (13) and (14) can be expressed in a general form as

$$\mathbf{p}_{m+1} = \mathbf{p}_m - \delta, \tag{19}$$

where $-\delta$ is the (vector) increment added to $\mathbf{p}_m$ in order to generate $\mathbf{p}_{m+1}$. Thus, for the steepest descent method we have

$$\delta_l = \gamma \beta_l . \tag{20}$$

It can be shown that $\boldsymbol{\alpha} = {}^1/_2\mathbf{D}$ in Equation (14), so that for the quadratic expansion method

$$\sum_{j=1}^{J} \alpha_{lj}\delta_l = \beta_l \,. \tag{21}$$

The Levenberg-Marquardt search combines these expressions by taking two further steps. First, based on dimensional arguments, the γ is chosen to be a vector whose l^{th} element is given by

$$\gamma_l = \frac{1}{\lambda\alpha_{ll}}$$

so that

$$\delta_l = \frac{1}{\lambda\alpha_{ll}}\beta_l \,. \tag{22}$$

In the above equations, λ is a non-dimensional parameter whose value we can freely vary depending on the search behaviour that we desire. The second step is to note that both Equations (21) and (22) can be written as a single expression

$$\sum_{j=1}^{J} \alpha'_{lj}\delta_l = \beta_l \tag{23}$$

where

$$\begin{aligned} \alpha'_{ll} &= \alpha_{ll}(1+\lambda) \\ \alpha'_{lj} &= \alpha_{lj} \quad \text{for}(j \neq l). \end{aligned} \tag{24}$$

For very large values of λ, α' is dominated by its diagonal elements so that Equation (23) approaches Equation (20). Conversely, as λ approaches zero, Equation (23) approaches Equation (21). Thus, simply by varying λ, we can smoothly vary the strategy for generating $\mathbf{p}_{m+1}$ between a steepest descent method and a quadratic expansion of χ^2 method. The former is most useful a long way from the χ^2 minimum, whereas the latter works best close to the minimum.

Press *et al.* (1986) describe the procedure for actually implementing an LM search as follows:

1. Generate an initial estimate of the parameters, $\mathbf{p}$, and evaluate $\chi^2(\mathbf{p})$.
2. Pick a starting value for λ, with 0.001 being recommended.
3. Evaluate the vector $\boldsymbol{\beta}$ and the matrix $\boldsymbol{\alpha}'$.
4. Solve Equation (23) for $\boldsymbol{\delta}$ by inverting $\boldsymbol{\alpha}'$.
5. If $\chi^2(\mathbf{p} - \delta) \geq \chi^2(\mathbf{p})$ increase λ by a factor of ten and go back to (3) .
6. If $\chi^2(\mathbf{p} - \delta) < \chi^2(\mathbf{p})$ decrease λ by a factor of ten and replace $\mathbf{p}$ with $\mathbf{p} - \delta$, and go back to (3) if the stopping criterion has not been satisfied.

7. Terminate the iteration when differences between successive χ^2 estimates are too small to be statistically meaningful (i.e. << 1).

8. Set $\lambda = 0$ and evaluate the co-variance matrix of standard errors, $\mathbf{C} = \alpha^{-1}$.

6.6.2 Application to the lidar analysis

The above description has described the Levenberg-Marquardt search in general terms. For the lidar problem, there are only two parameters that we need to find by searching, i.e. T and λ_0 Thus, β is a two-element vector and α' is a 2 x 2 matrix. Because these each have a small number of elements, solving Equation (23) is not computationally expensive.

Application of the LM method to any specific problem, including that of the lidar analysis, requires deriving expressions for $\partial\chi^2/\partial p_j$ so that β and α' can be evaluated. Now Equations (16) and (18) indicate that we do not need to evaluate $\partial\chi^2/\partial p_j$ directly; we can adequately approximate these terms provided we know the derivatives of the basis functions, $\partial s^n/\partial p_j$. These latter expressions are more simply obtained, as is done below.

6.6.3 The partial derivative with respect to Doppler shift

Expanding Equation (2) to explicitly show all its terms yields

$$s_n = a_0 + a_1\left[I(\lambda - \lambda_0)\right]_n + a_2\left[I * \sum_{k=0}^{K-1} q_k \exp -\left(\frac{\lambda - \lambda_0}{w_k(T)}\right)^2\right]_n,$$

from which we obtain the result that

$$\frac{\partial s_n}{\partial \chi_0} = a_1\left[\frac{\partial}{\partial \lambda_0} I(\lambda - \lambda_0)\right]_n + a_2\left[I * \sum_{k=0}^{K-1} \frac{2(\lambda - \lambda_0)}{w_k^2(T)} q_k \exp -\left(\frac{\lambda - \lambda_0}{w_k(T)}\right)^2\right]_n. \quad (25)$$

Notice that the derivative is explicitly evaluated in the second (molecular scattering) term of Equation (25) above, but not in the first (aerosol scattering) term. For the second term, we have evaluated the derivative using the property that

$$\frac{\partial}{\partial u}\left[I(\lambda) * f(\lambda, u)\right] = I(\lambda) * \left[\frac{\partial}{\partial u} f(\lambda, u)\right]. \quad (26)$$

Now because $I(\lambda - \lambda_0) = I(\lambda) * \delta(\lambda - \lambda_0)$, we can indeed rewrite

$$\frac{\partial}{\partial \lambda_0} I(\lambda - \lambda_0)$$

in the same form as the left-hand side of Equation (26) above. However, because we can neither represent (much less differentiate) a discretely-sampled delta function, we cannot use the right-hand side of Equation (26) to evaluate the aerosol derivative in practice. Further,

because $I(\lambda)$ is an experimentally determined function (i.e. it is not analytically described), we can never express

$$\frac{\partial}{\partial \lambda_0} I(\lambda - \lambda_0)$$

analytically – we must compute it numerically. The computer program actually does this by applying the Fourier derivative theorem to the discrete Fourier transform $\{I_n\}$. ($\{I_n\}$ is the discrete representation of $I(\lambda)$).

6.6.4 The partial derivative with respect to temperature

As before, we expand Equation (1) to show all its terms. However, this time we must explicitly include the expression that relates $^1/_e$ Doppler half width of the molecular backscatter (in measurement units) to kinetic temperature, i.e.

$$w(T) = \xi \lambda_0 \sqrt{\frac{2kT}{Mc^2}}$$

where ξ the coefficient needed to convert the wavelength units (of λ_0) to 'channels' in our discretely sampled spectra. Our expression thus becomes

$$s_n = a_0 + a_1 \left[I(\lambda - \lambda_0) \right]_n + a_2 \left[I^* \sum_{k=0}^{K-1} q_k \exp - \left(\frac{\lambda - \lambda_0}{\xi \lambda_0 \sqrt{\frac{2kT}{M_k c^2}}} \right)^2 \right]_n .$$

Differentiating with respect to T gives

$$\frac{\partial s_n}{\partial T} = a_2 \left[I^* \sum_{k=0}^{K-1} \frac{M_k c^2}{2kT^2} \frac{(\lambda - \lambda_0)^2}{\xi^2 \lambda_0^2} q_k \exp - \left(\frac{\lambda - \lambda_0}{\xi \lambda_0 \sqrt{\frac{2kT}{M_k c^2}}} \right)^2 \right]_n . \qquad (27)$$

Equation (27) completes the information needed to calculate the vector β and the α' used in the LM search.

6.7 The complete lidar analysis scheme

We now consider how the methods above can be combined to form a complete lidar analysis scheme. We have constructed the analysis in two parts: one to solve analytically for parameters appearing linearly in our model s_n, and one to solve for the non-linear parameters. We perform these two parts iteractively. The procedure is:

1. Generate initial guesses for T and λ_0.
2. Generate new estimates for a_0, a_1 and a_3 using the SVD method and the T and λ_0 estimates.

3. Generate new estimates for T and λ_0 using an LM search and the current a_0, a_1 and a_3 estimates.

4. Evaluate χ^2 using the updated parameters.

5. Go back to (2) until successive χ^2 estimates are not distinguishable with statistical significance.

One slight complication is that the LM method itself is an iterative procedure. Thus, the analysis could either leave the linear parameters constant throughout each complete LM search, or alternatively, it could re-calculate the linear parameters between every step in the LM search. In practice, the former approach is faster.

6.8 Estimating the parameter variances

Both the SVD and the LM portions of the complete solution can return estimates of the variances of the parameters. However, while it is simple to independently compute values of a subset of the model parameters (given values for each of the remaining parameters), it is not simple to independently compute the variances of the subset (even given values and variances of the remaining parameters). Fortunately, standard non-linear least-squares fitting techniques (e.g. Bevington, 1969; Press *et al.*, 1986) provide a generalised method for computing the variances of all the model parameters simultaneously, as follows.

Let us denote the complete set of L variable parameters in our model as $\{v_l\}$ where l varies from 0 to $L - 1$. (Using the notation of the previous sections, $\{v\}$ would be formed from the union of the sets $\{a\}$ and $\{p\}$ and $L = 5$ if we fit all parameters.) We construct a complete L x L curvature matrix, $\mathbf{c}$, encompassing all parameters using

$$c_{jk} = \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} \frac{\partial s_n}{\partial v_j} \frac{\partial s_n}{\partial v_k} . \tag{28}$$

The parameter variances are then obtained directly from the diagonal elements $\mathbf{c}^I$. That is, if we write $\varepsilon = \mathbf{c}^I$, then

$$\sigma^2(v_l) = \varepsilon_{ll} . \tag{29}$$

6.9 Performance evaluation

6.9.1 Simulating spectra

To examine the performance of the method, we apply it to simulated Fabry-Perot spectra, generated numerically from precisely known values of the five model parameters. By applying the method to large sets of such spectra, we can examine the speed and stability of convergence, as well as the statistical distributions of the fitted parameters. In particular, we need to know if the means of the parameter estimates match the values used to generate the spectra. Further, we will compute the variances of the distributions of parameter estimates over the whole set of simulated spectra, and see if these are consistent with the parameter

variances returned by the fit routine after fitting an individual spectrum. The basic characteristics used to simulate the experiment were:

Laser wavelength	532 nm
Etalon gap	15.8 mm
FPS instrumental finesse	19.7
Number of spectral channels	128
Doppler shift by one channel	~ 40 m s^{-1}
Atmospheric composition	75% N_2, 25% O_2
Doppler width	250 Kelvins
Doppler shift	20 channels
Molecular peak height	50 counts
Aerosol peak height	15 counts
Spectral background	45 counts/channel

In simulating the instrument profile, an attempt was made to include several of the 'less desirable' characteristics that frequently occur in real instruments. Thus, the simulated instrument profile was generated by summing three Gaussian functions, and then convolving these with an asymmetric (right-angle triangle) shaped function. For the Gaussians, the dominant one was chosen to be very narrow (corresponding to an equivalent N_2 temperature of eight Kelvins). The remaining two Gaussians heights were only 30% and 10% that of the main one, but their spectral widths were five and ten times greater, respectively. The full-width at half height (FWHH) of the triangle function was set to be twice that of the main Gaussian's FWHH. The instrument function

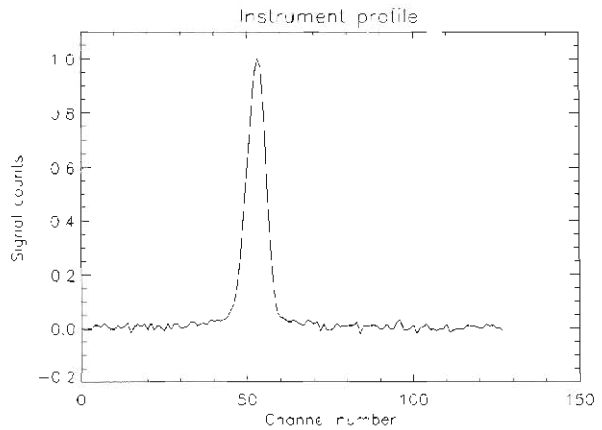


Figure 1. *The simulated instrument profile used to test the analysis procedure. This is the version actually supplied to the analysis routine, including the noise term.*

was deliberately displaced slightly off-center of the spectrum; its centroid occurs at 40% of the scan range. A noise-free version of this instrument function was used to generate the simulated sky return spectra. However, the version supplied to the fitting routine had a small noise term added. This was to simulate the real experiment, in which the instrument function can only be obtained by measurement – a process that inevitably introduces noise. Figure 1 illustrates the simulated instrument profile supplied to the program, including measurement noise.

6.9.2 Some example fits

We now present three examples of the results of fitting to individual spectra. Figure 2 presents an example of a simulated sky return spectrum, with the fitted function superimposed, after 0, 3 and 12 iterations of the procedure. To test the fitting routine strenuously, the initial estimates of peak position and temperature were allowed to vary randomly over a much wider range than would occur in practice. In Figure 2 the initial peak position estimate was clearly to high. Nevertheless, convergence was good, as is illustrated by Figure 3 which depicts the history of the estimates for each parameter after each iteration. Indeed, tests have shown that the procedure can converge from initial guesses very much worse than shown here. The only (minor) difficulty occurs if the initial temperature is absurdly low and the initial position is several peak widths away from its correct location. In this case there is little overlap between peaks in the data and the model, yielding a poor indication of which way the position and temperature parameters need adjusting. However, the program can test for this condition, and can respond effectively.

The program actually terminated after a total of 21 iterations, when the reduced χ^2 value was 1.20. (21 iterations is more than usual; the average number of iterations at termination is only 13.) Included among the diagnostic information returned by the program is the sky spectrum signal/noise ratio. This is computed from the power spectrum of the sky profile, and is defined as the power spectrum's fundamental component divided by the average value of it's high-frequency, noise-dominated components. The signal/noise ratio was 748 for the sky spectrum shown in Figure 2 (top panel). The final parameter estimates and their corresponding uncertainties were:

	Value	Uncertainty
Temperature	243.6 Kelvins	26.3 Kelvins
Doppler shift	19.9 channels	0.24 channels, or 9.6 ms-1
Molecular	48.9 counts	2.9 counts
Aerosol	16.3 counts	4.1 counts
Background	46.4 counts/channel	0.58 counts/channel

The sequence of χ^2 values after each iteration is shown in Figure 4.

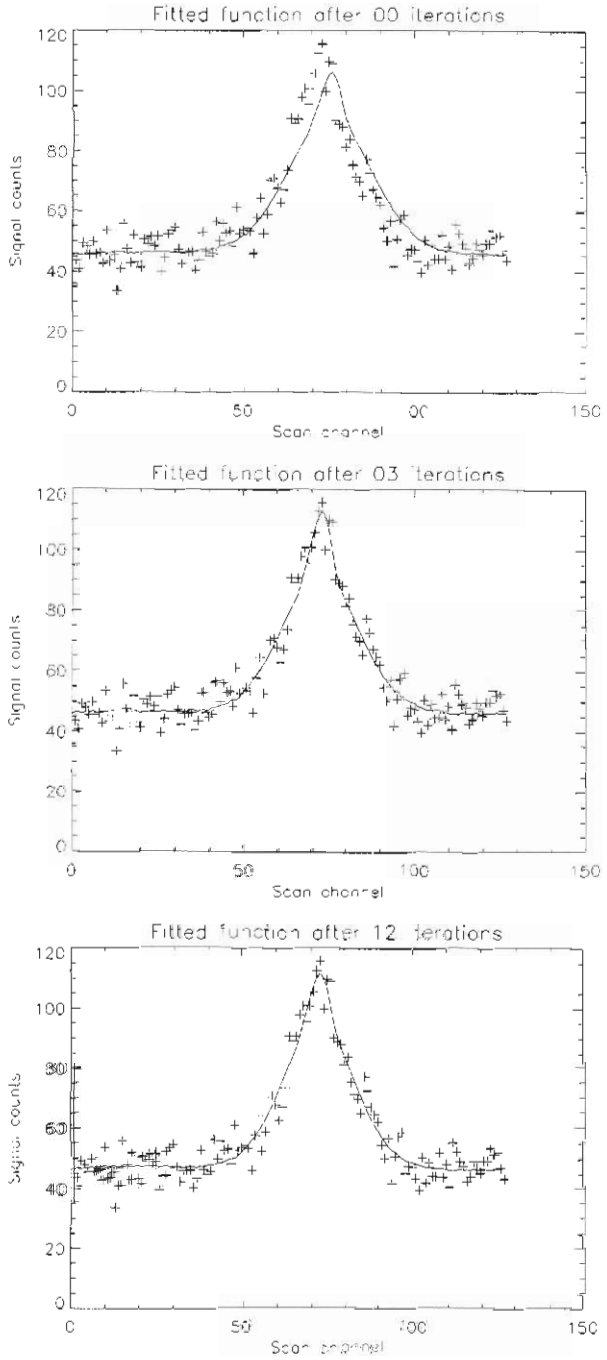


Figure 2. *These three panels show a simulated Fabry-Perot spectrum (including noise) and the fitted function after 0, 3 and 12 iterations respectively. The function at zero iterations uses intensities (a_0, a_1 and a_2) calculated from the initial guesses for temperature and Doppler shift. The change to the fitted function between 3 and 12 iterations was minor, but discernable.*

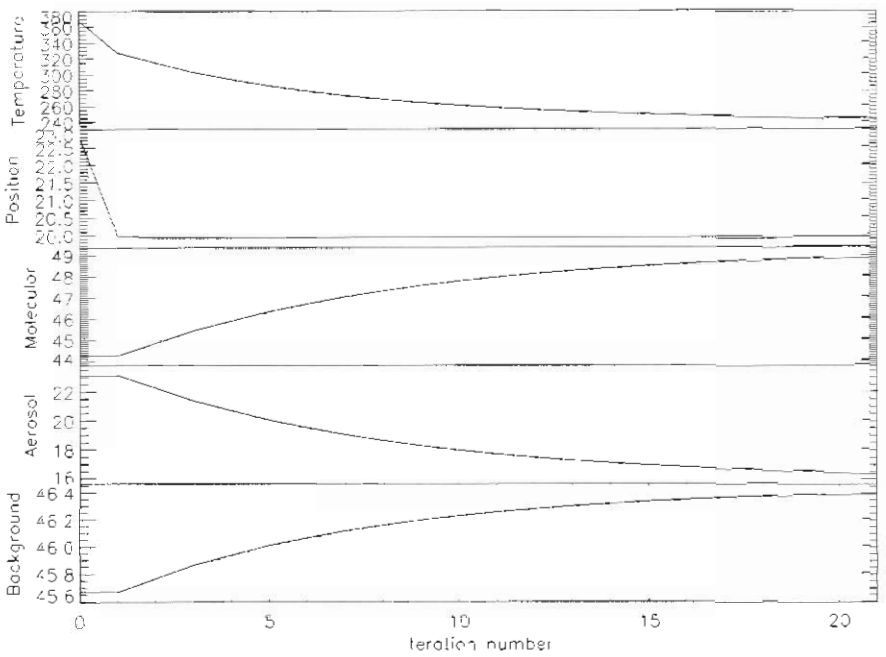


Figure 3. *These five panels show the sequence of estimates for each of the five model parameters, as functions of the analysis iteration number. Each parameter converged stably. Note the Y-axis label 'Position' for the second panel down is actually synonymous with 'Doppler shift'.*

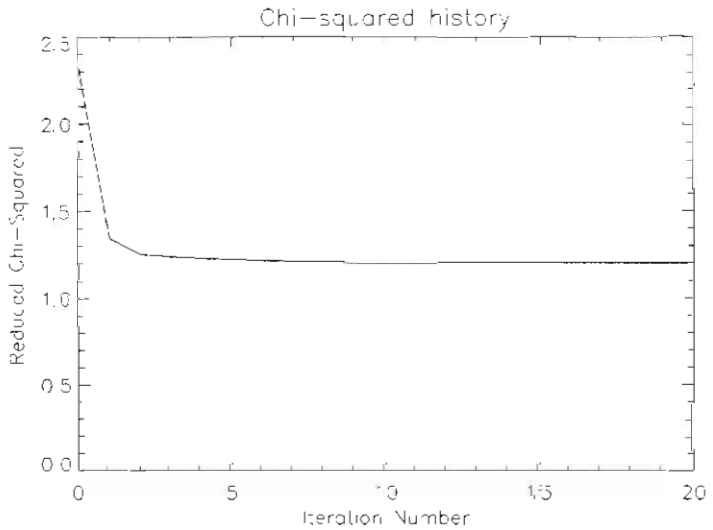


Figure 4. *The sequence of reduced χ^2 values as a function of iteration number, for the fit to the sky spectrum shown in Figure 2.*

Figure 5 shows an example of the fit results for a spectrum with a low signal/ratio, i.e. only 91. The program terminated after a total of 13 iterations, when the reduced χ^2 value was 0.84. The final parameter estimates and their corresponding uncertainties were:

	Value	Uncertainty
Temperature	272.0 Kelvins	90.0 Kelvins
Doppler shift	19.7 channels	0.70 channels, or 28 m s^{-1}
Molecular	48.3 counts	8.6 counts
Aerosol	21.4 counts	12.4 counts
Background	43.7 counts/channel	1.8 counts/channel

Figure 6 shows an example of the fit results for a spectrum with a high signal/ratio, i.e. 2194. The program terminated after a total of 22 iterations, when the reduced χ^2 value was 0.91. The final parameter estimates and their corresponding uncertainties were:

	Value	Uncertainty
Temperature	257.3 Kelvins	16.4 Kelvins
Doppler shift	19.98 channels	0.14 channels, or 5.6 m s^{-1}
Molecular	48.3 counts	1.7 counts
Aerosol	16.3 counts	2.4 counts
Background	45.5 counts/channel	0.34 counts/channel

6.9.3 Distributions of the fit results

Having seen how the method behaves in a few selected examples, we now consider the statistical distributions of fit results for a large (1500 element) set of simulated spectra. The amplitude of the noise term added to each spectrum was varied randomly, so that the set of spectra spanned a range of signal/noise ratios. Figure 7 presents scatter diagrams of the values returned for each of the five model parameters, plotted against signal/noise ratio. It is apparent that the parameter estimates cluster about their correct values for all five parameters and at all noise levels.

As expected, the parameter estimates become more widely scattered (i.e., the parameter variances increase) at higher noise levels. While generating the results in Figure 7, the method took an average 13 iterations and 0.45 seconds of computer time per spectrum. Thus, it took around 12 minutes to process all 1500 spectra.

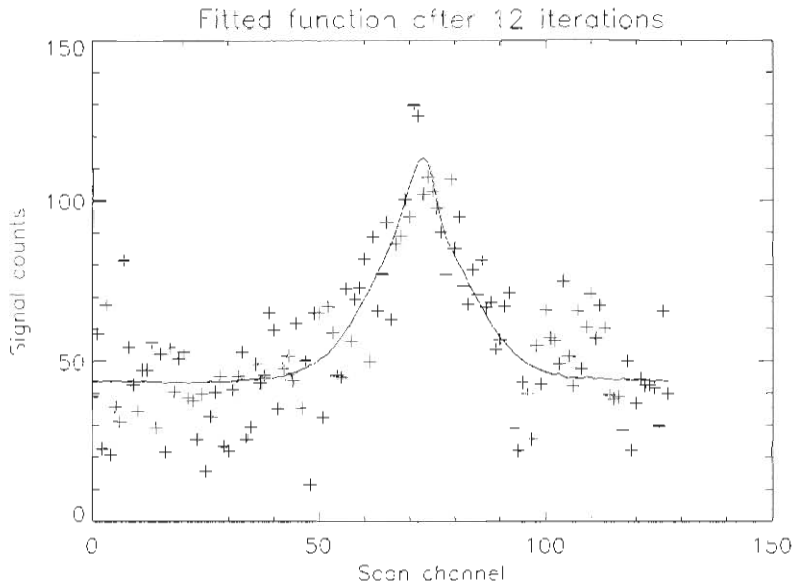


Figure 5. *A low signal/noise ratio spectrum (i.e. signal/noise = 91) and the corresponding fitted function.*

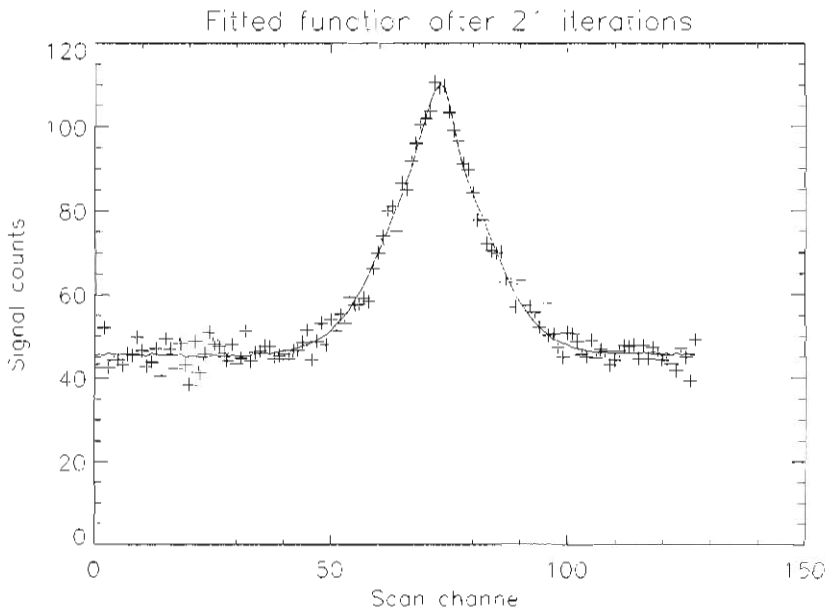


Figure 6. *A high signal/noise ratio spectrum (i.e. signal/noise = 2194) and the corresponding fitted function.*

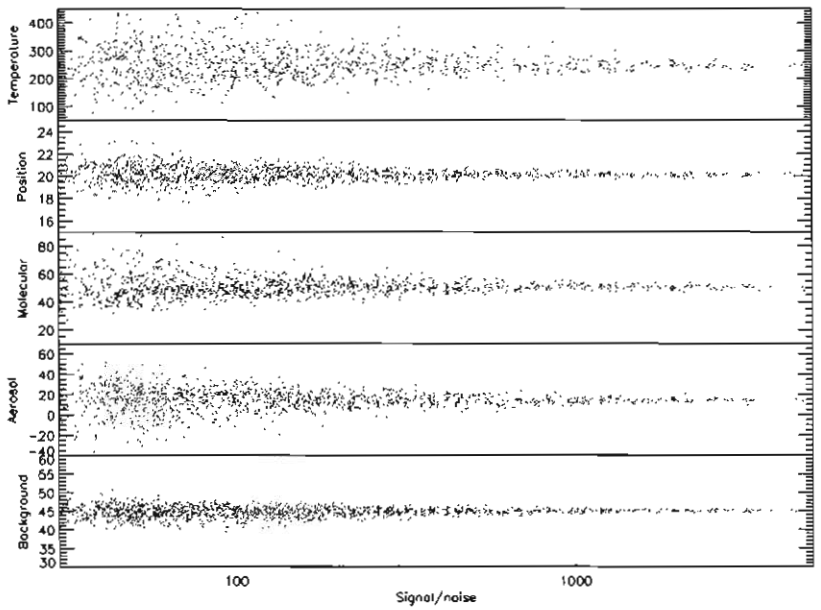


Figure 7. *Scatter plots of parameter estimates versus signal/noise ratio, for a set of 1500 numerically generated test spectra.*

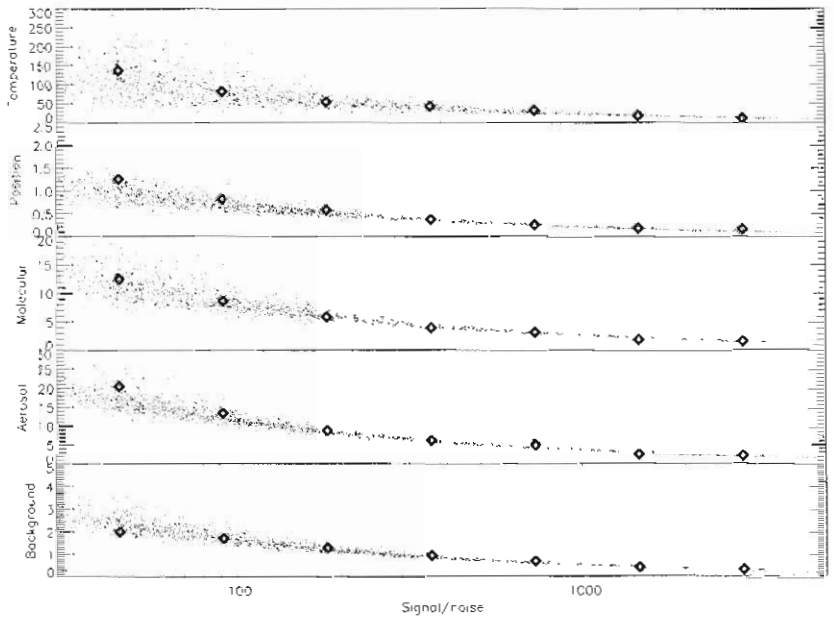


Figure 8. *Scatter plots of parameter standard deviations returned by the analysis program, plotted against spectral signal/noise ratio (small points). Superimposed as heavy diamonds are the actual standard deviations of the parameter estimates obtained in seven subsets of the spectra, spanning seven contiguous intervals of signal/noise ratio.*

Figure 8 presents scatter diagrams of the parameter standard deviations (i.e. $\sqrt{\sigma^2}$)returned by the fit, plotted against signal/noise ratio. To test if the variances calculated in Equation (29) appear reasonable, we compared these values to the distributions of parameter estimates within seven subsets the entire 1500 spectra. That is, we divided the spectra into seven subsets, spanning seven contiguous intervals of signal/noise ratio. Within each of these subsets we computed the variances of the parameter estimates. The square root of these values represents the actual standard deviations that were obtained within that subset; in Figure 8 these values are superimposed as heavy black diamonds onto the scatter plots of individual standard deviation estimates. The variances (and hence standard deviations) returned from fitting an individual spectrum generally appear to be consistent with actual variances over a large set of trials. There is perhaps a suggestion that the uncertainties in Doppler shift and aerosol scattering intensity may be slightly underestimated for signal/noise ratios below ~ 100. However, this is an inexact test of the routine, and the agreements are close. Further, spectra with signal/noise ratios this low are unlikely to be of much use anyway. Overall, refinement of the error estimation method does not seem justified for now.

6.9.4 The effects of fitting fewer parameters

As mentioned earlier, the analysis program has been written to allow one or more of the model parameters to be held 'frozen' at some externally-supplied initial value. This facility was included as it is anticipated that several of the parameters could potentially be estimated independently by other means, for a typical lidar experiment. It is instructive to see how inclusion of such *a-priori* knowledge can improve the fit results for the remaining parameters. Figure 9 shows the parameter standard deviations for a second set of 1500 spectra, statistically equivalent to those used to generate Figure 8. However, in this case, the fits were constrained to only search for temperature, Doppler shift, and molecular backscatter intensity. The aerosol and background intensities were held fixed, at their correct values.

The uncertainties in temperature and molecular scattering intensity were dramatically reduced by applying *a-priori* knowledge of the background and aerosol intensities. A discernible, but less dramatic, improvement was also seen for the Doppler shift parameter. The computational load was also reduced. On average, the routine converged after only nine iterations and took 0.32 seconds per spectrum.

6.9.5 The effects reduced instrumental finesse

The simulations presented so far have used an instrument function finesse ~ 20, which is toward the upper limit of what is likely to be achievable for a practical FPS. We now consider the results of a simulation using a broader instrument function, with a finesse of ~ 9, as depicted in Figure 10. This instrument function was generated using the same procedure as previously, except that the main Gaussian's equivalent N_2 temperature was specified to be 24 Kelvins (rather than

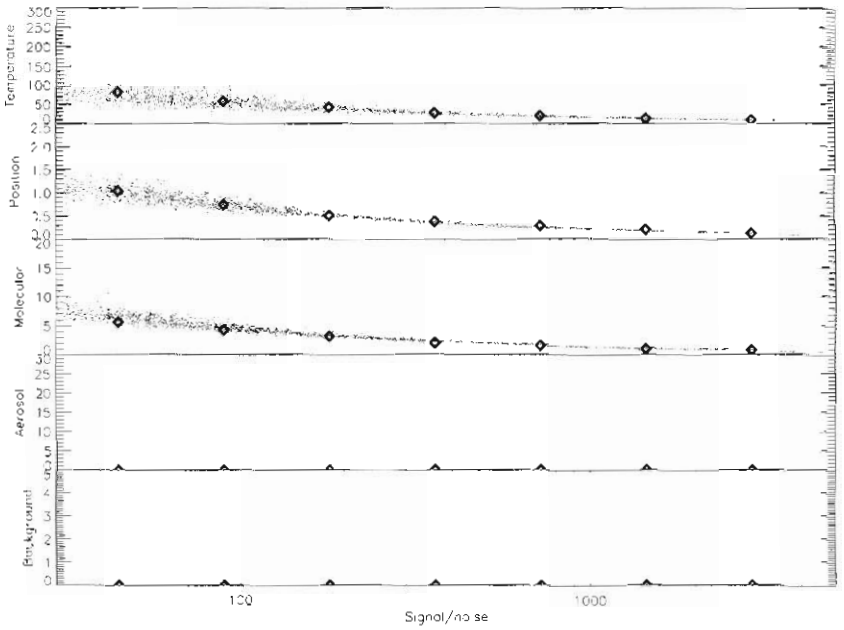


Figure 9. *As before, this figure shows scatter plots of parameter standard deviations, both estimated for individual spectra and calculated from the distributions. In this case, only the top three parameters were fitted. The bottom two were held fixed at their correct values throughout the fit.*

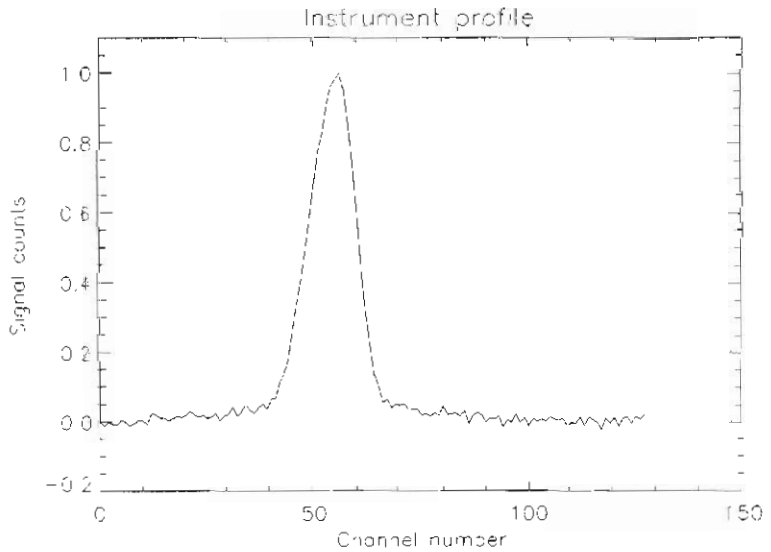


Figure 10. *This figure depicts an instrument profile approximately twice the width of that used previously. Simulations using this profile were run to examine how reduced instrument finesse would degrade the recovery of model parameters.*

eight Kelvins). The widths of the other contributing functions were scaled relative to that of the main Gaussian, so all component functions were similarly broadened, yielding a final instrument function width roughly twice that used previously. (Note, however, that at 14.3 channels, the FWHH of this function is still significantly less that the FWHH of 22.9 channels for scattering from N_2 at 250 Kelvins.)

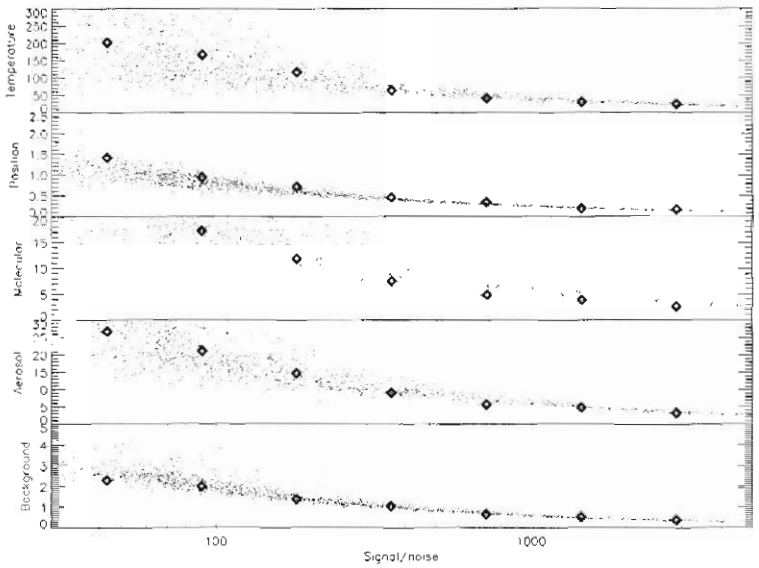


Figure 11. *This figure shows scatter plots of parameter standard deviations from analysis of spectra generated using the broadened instrument function. Molecular and aerosol scattering intensities are the parameters most degraded by the reduced finesse, followed by temperature.*

Figure 11 shows the parameter standard deviations for a another set of 1500 spectra, in this case generated using the broadened instrument function depicted in Figure 10. As can be seen, the parameters worst effected were molecular and aerosol scattering intensities, followed by temperature. Some degradations also occurred for the background intensity and Doppler shift, but these were less significant. It would seem that a high finesse instrument is most critical for distinguishing between aerosol and molecular scattering, and for measuring temperature. If wind is the only quantity of interest, there is less incentive to achieve high finesse in the Fabry-Perot system.

6.10 Conclusions

A new method has been presented for analysing Doppler spectra from an atmospheric lidar experiment. The method allows for the effects of both aerosol and multiple-species molecular scattering, a continuum background, and instrumental broadening of the recorded spectra.

The analysis proceeds by least-squares fitting a five-parameter model function to the recorded spectra. Values for three of these parameters are computed analytically, and two by numerical searching. One or more parameters can be held fixed at externally-supplied values, while the remainder are computed accordingly. Applying the method to simulated spectra indicated that it converges stably over a wide range of signal/noise ratios, and even given very poor initial guesses for the two search parameters. Parameter estimates from sets of 1500 trial spectra indicated that the means of the parameter estimates closely matched the values used to generate the model spectra. Further, the variances of the sets of results were, in most cases, consistent with the variances returned by individual fits. There was perhaps a suggestion that the variances in Doppler shift and aerosol scattering intensity were slightly underestimated for signal/noise ratios ~ 100. However, the discrepancy did not appear large enough to warrant further refinement of the error estimation for now.

The simulations indicated that significant improvement in the accuracy of estimation of some parameters can be obtained by minimising the number of parameters solved simultaneously, and by making the instrument function as narrow as possible.

Finally, simulations like these have proved useful not just for testing the analysis procedure. They have also yielded useful guidance for the design of the lidar experiment itself. For example, it is apparent that there is less incentive to achieve a narrow instrument function if Doppler shift is the only quantity needed with high precision.

6.11 Acknowledgments

This work was supported by the Australian Antarctic Division. The author visited the Antarctic Division's Atmospheric and Space Physics section between July and September 1998, at the kind invitation of Dr Ray Morris.

References

Bevington, P.R. (1969). *Data reduction and error analysis for the physical sciences.* McGraw-Hill Book Company, New York.

Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1986). *Numerical recipes: the art of scientific computing.* Cambridge University Press.

7. A MODERN, INITIAL ATTEMPT TO DETERMINE TEMPERATURES FROM A ROTATING-FILTER PHOTOMETER MONITORING THE MAJOR P-BRANCH LINES OF THE OH(6-2) BAND AT DAVIS, ANTARCTICA

P.F.B. Williams[1], G.B. Burns[1], M. Lambert[1] and W.J.R. French[1,2]

[1] Atmospheric and Space Physics
Australian Antarctic Division
Kingston Tasmania 7050
(email: gary.burns@aad.gov.au)

[2] Institute of Antarctic and Southern Ocean Studies
University of Tasmania
Hobart Tasmania 7000

Abstract

Photometric intensity measurements of the OH(6-2) band P1(2), P1(3), P1(4) and P1(5) emissions were collected at a four minute repetition rate at Davis station, Antarctica during 1994. These data are being examined to determine the magnitude, percentage occurrence and duration of temperature fluctuations in the 15 minute to two hour period range. A report on the progress of the work is presented.

7.1 Introduction

Hydroxyl (OH) emissions originate from a layer near 87 km. The atmospheric emission has a number of bands which are defined by vibrational state transitions. By comparing the relative intensity of two or more lines within one branch of a band the rotational temperature of the molecule can be determined. If the OH molecules are in thermal equilibrium with the atmosphere, the rotational temperature can be interpreted as the kinetic atmospheric temperature of the near-mesopause, emission region. For a more extensive introduction to the OH(6-2) band, readers are referred to Greet *et al.* (1998).

7.2 Data selection

Photometric intensity measurements of the OH(6-2) band P1(2), P1(3), P1(4) and P1(5) emissions were collected at a 4-minute repetition rate at Davis station, Antarctica during 1994 using a zenith-oriented, rotating-filter photometer with a 6° field-of-view.

Prime data intervals are when the moon was below the horizon, the Sun was more than 18° below the horizon (astronomical twilight) and local visual determinations (assisted by a video airglow imaging system) of

clear sky. Figure 1 shows the data intervals that passed these criteria, amounting to some 368 hours over 51 separate nights.

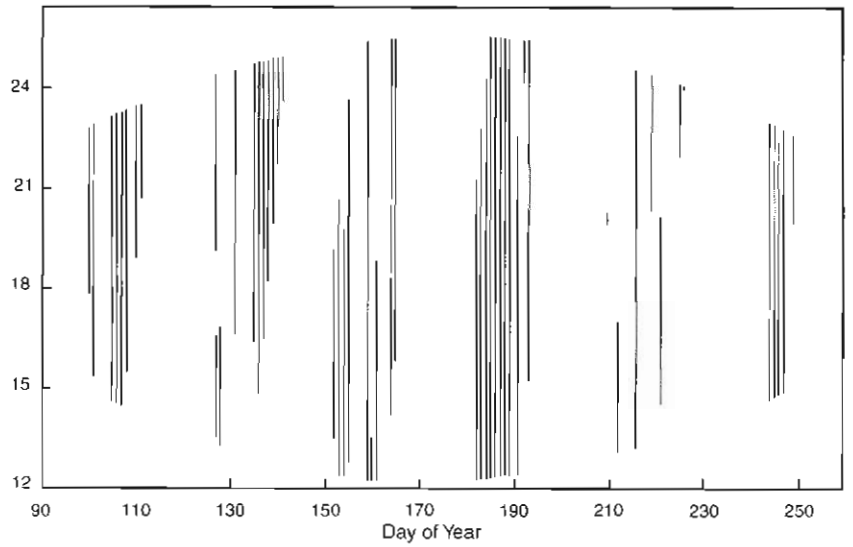


Figure 1. *Davis 1994 photometer observation times for the Sun more than 18° below the horizon, no moon and no cloud.*

7.3 Temperature determinations in the OH(6-2) band

Figure 2 shows the large scale features of the OH(6-2) band. Figure 3 shows the background features in the vicinity of the major lines identified by an associated study using a Czerny-Turner spectrometer. The auroral features noted in Figure 3 result from second order N2+(1NG) band emissions and will not effect the photometric measurements.

Figure 4 shows a typical measurement of the filter profiles taken during 1994 and 'best-fit' Gaussian profiles. The spectral width of the filters (full-width-at-half-maximum of the order of 0.7 nm for the major lines) means contributions from a number of near-by features must be allowed for in deriving the measured intensities of the major lines. The minor features include sub-satellite (cross-branch) features that have been identified, and magnitudes determined, by the associated study of Czerny-Turner profiles of the OH(6-2) band region. We make allowance for Q1(5) under P1(2) and Q1(6), QR12(2+3) and QR12(1) under P1(3).

Four measurements of the filter profiles (three only for P1(3)) were made during 1994. Gaussian curves were best-fitted to the measured profiles and the effective filter width (the width of the filter scaled to 100% transmission – required for background contribution estimations) and the percentage transmission at major lines and background features estimated. Table 1 shows the percentage transmissions estimated. Note: Hydroxyl emissions are doublets (denoted by e and f; central transmission estimates are denoted by c).

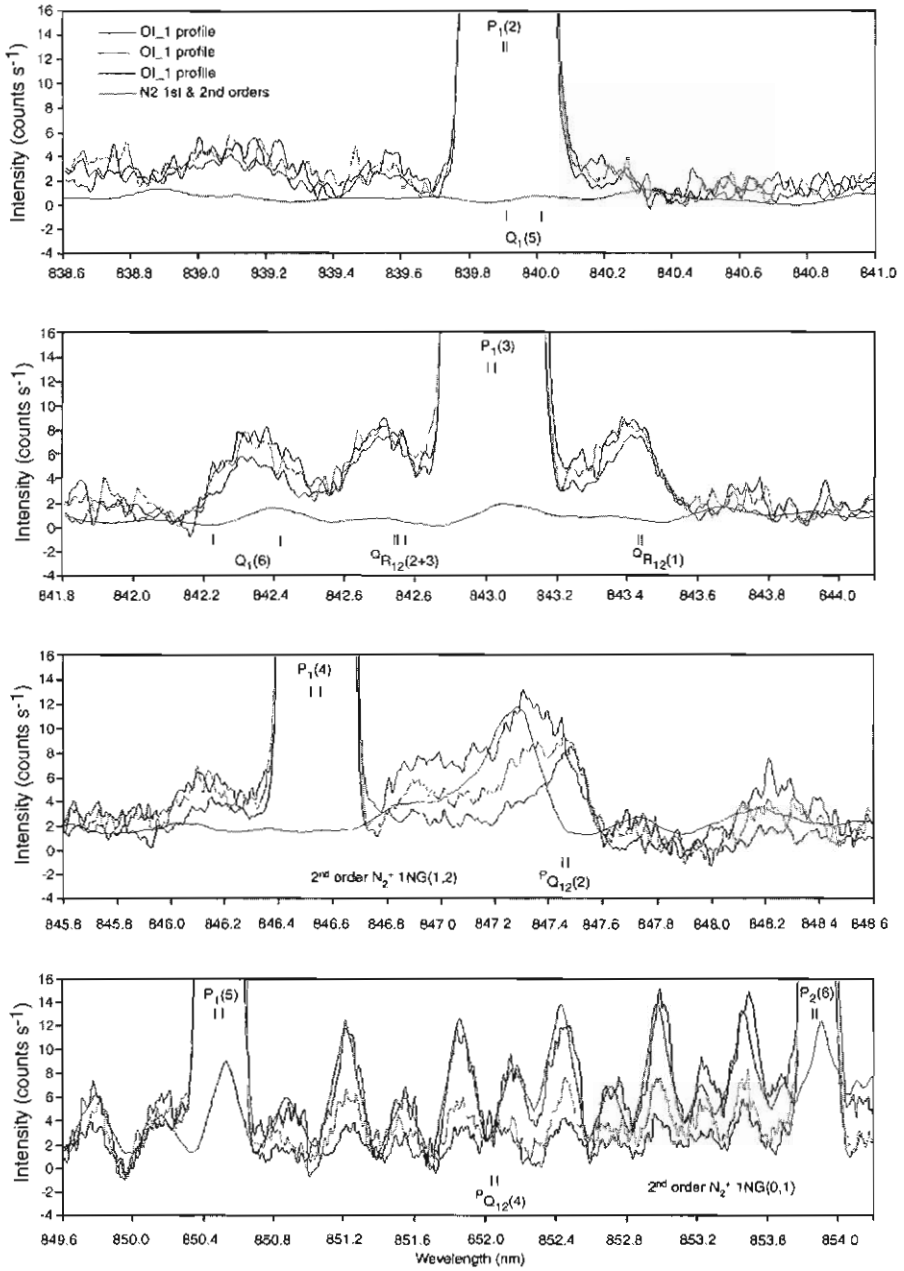


Figure 2. *The major emission features of the OH(6-2) band.*

Straight lines were fitted to the measured transmissions to estimate the transmissions on any day. Errors for these fits have been determined. Table 2 lists these linear fits and errors. For the P1(3) filter transmissions, average slopes from the other fits were assumed due to the reduced number (3) of measured profiles.

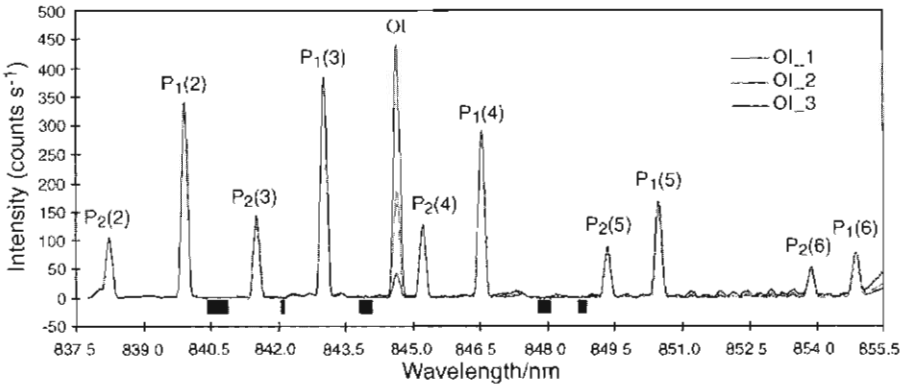


Figure 3. *Averaged profiles at three levels of auroral activity showing the auroral dependence of background features in the vicinity of the lines. The auroral features predominantly result from second-order fold-in which will not effect the photometer measurements.*

The background filter used was so broad (fwhm of the order 1.3 nm, see Figure 4) that a 'not insignificant' contribution to the total counts comes from the P1(4) and P2(5) lines (see Tables 1 and 2). We have determined a background value by assuming that it is equivalent (in counts nm^{-1} s^{-1}) across the four major lines and determining a background value for each filter rotation by minimising the standard deviation in the temperatures derived from each of the six major line ratios [P1(2)/P1(3), P1(2)/P1(4), P1(2)/P1(5), P1(3)/P1(4), P1(3)/P1(5) and P1(4)/P1(5)].

Figures 5 and 6 show average temperatures and temperatures derived from each of the six ratios for two data intervals. In both cases, long period fluctuations of magnitude 20 K are apparent. The corrected line intensities are also shown, along with the background value determined by minimising the standard deviation in the temperatures derived from each of the six major line ratios. For the long period fluctuations shown, the temperature variations appear to lead the intensity fluctuations by the order of 15 to 30 minutes.

The examples shown in Figures 5 and 6 indicate that the backgrounds derived vary, but they show trends not inconsistent with a varying stellar contribution to a 6° field-of-view photometer.

Table 1. *The percentage transmission of various OH lines through the photometer filters for measurements made on the days indicated.*

OH line	λ (nm)	70	122	177	209	: day	
P1(2)e	839.91	39.68	39 73	35.37	36.87		
P1(2)c	839.92	39.74	39.75	35.40	36.88		
P1(2)f	839.92	39.79	39.77	35.42	38.88		Filter 1
Q1(5)f	839.91	39.71	39.74	35.38	36.87		
Q1(5)c	839.96	39.61	39.29	35.12	36.34		
Q1(5)e	840.02	38.02	37.38	33.62	34.52		
P1(3)e	843.01		29.11	24.96	26.47		
P1(3)c	843.02		28.95	24.73	26.25		
P1(3)f	843.03		28.76	24.48	25.99		
Q1(6)f	842.23		1.73	2.34	2.27		Filter 2
Q1(6)c	842.32		3.47	4.37	4.30		
Q1(6)e	842.42		6.30	7.43	7.42		
QR12(1)	843.44		8.68	6.53	7.06		
QR12(2)	842.75		23.35	22.58	23.47		
QR12(3)	842.77		24.56	23.46	24.44		
P1(4)e	846.52	42.01	40.65	37.73	38.92		Filter 3
P1(4)c	846.54	41.95	40.40	37.36	38.53		
P1(4)f	846.55	41.77	40.03	36.93	38.04		
P1(5)e	850.46	39.72	37.38	34.87	35.51		Filter 4
P1(5)c	850.48	39.56	36.89	34.38	34.92		
P1(5)f	850.51	39 18	36.23	33.74	34.16		
P1(4)e	846.52	1.53	1.97	2.55	2.24		Filter 5
P1(4)c	846.54	1.65	2.12	2.73	2.41		
P1(4)f	846.55	1.78	2.27	2.92	2.58		
P2(5)f	849.33	2.76	2.45	2.46	2.24		
P2(5)c	849.34	2.69	2.38	2.40	2.18		
P2(5)e	849.34	2.61	2.31	2.33	2.12		
PQ12(2)fe	847.45	38.39	40.15	40.99	40.22		
PQ12(2)c	847.46	39.08	40.80	41.58	40.83		
PQ12(2)ef	847.47	39.77	41.45	42.17	41.44		

Table 2. *The linear, annual trends of the transmission of the OH lines through the photometer filters.*

y	m	b	y error	% y error
width	-0.0001	0.277	0.008	3.1
P1(2)e	-0.0029	4.204	0.155	4.1
P1(2)f	-0.0029	4.218	0.153	4.0
Q1(5)f	-0.0029	4.207	0.155	4.1
Q1(5)e	-0.0031	4.042	0.116	3.2
width	-0.0001	0.240	0.008	3.8
P1(3)e	-0.0030	3.191	0.144	5.3
P1(3)f	-0.0033	3.205	0.145	5.4
Q1(6)f	0.0007	0.098	0.022	11.0
Q1(6)e	0.0014	0.474	0.034	4.9
QR12(1)	-0.0021	1.096	0.091	11.6
QR12(2)	0.0000	2.318	0.068	2.9
QR12(3)	-0.0003	2.472	0.083	3.4
width	-0.0001	0.309	0.003	0.9
P1(4)e	-0.0027	4.380	0.107	2.6
P1(4)f	-0.0032	4.383	0.109	2.8
width	-0.0002	0.304	0.003	1.0
P1(5)e	-0.0034	4.173	0.092	2.5
P1(5)f	-0.0039	4.143	0.096	2.7
width	-0.0003	0.874	0.002	2.7
P1(4)e	0.0006	0.118	0.027	12.6
P1(4)f	0.0007	0.139	0.029	12.1
P2(5)f	-0.0003	0.295	0.011	4.3
P2(5)e	-0.0003	0.279	0.010	4.4
PQ12(2)fe	0.0015	3.784	0.080	2.0
PQ12(2)ef	0.0013	3.930	0.077	1.9

y = mx+b **x = day number for 1994**

y = effective filter width (in nm) at 100% transmission or % transmission of tabulated lines

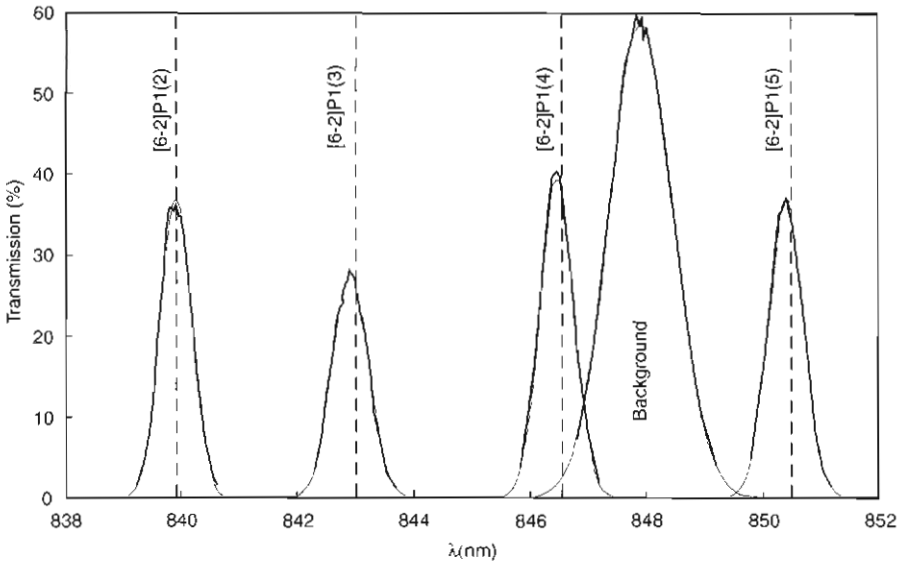


Figure 4. *Filter profile measurements and associated Gaussian fits.*

7.4 Possible analysis improvements

Recent work on the Czerny-Turner profiles have shown that the assumption of constant background can be improved on. A variation of up to 20% in this value across the OH(6-2) band appears likely. When appropriate factors are determined, the correction will be incorporated.

The method of determining the background by minimising the standard deviation in the temperatures derived from the six possible ratios is subject to chance variation. An alternative method is to determine an average linear relationship between the value determined from the background filter and the 'minimised variation' determination, and then to utilise that fraction of the background filter measurement.

It is possible to make a distinction between the uncertainty in the absolute temperature determination and the uncertainty in the temperature fluctuations. The uncertainty in the percentage transmission of the filters directly effects the absolute temperature determination, but has only minor influence on temperature variations. The determination of temperature fluctuations is limited primarily by photon counting statistics.

The effect of uncertainties in the filter transmissions on the magnitude of determined temperature fluctuations can be individually determined, and the combined effects estimated.

It is possible to determine temperature variations (and associated uncertainties) for individual ratios, and in a weighted manner for a combined estimation.

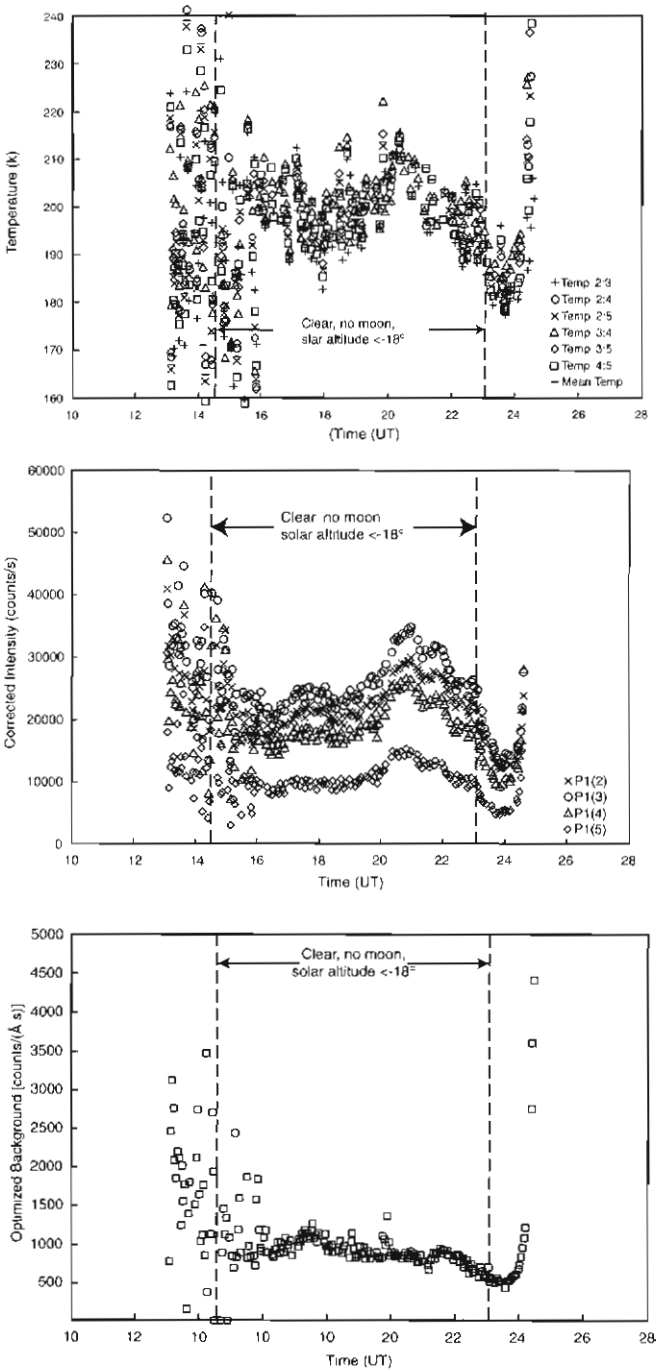


Figure 5. *Davis, 1994: 1300 UT–Day 107 to 0100 UT – Day 108*
(a) Temperatures derived from individual ratios, and their mean.
(b) Intensities of the major P-branch lines.
(c) Background, calculated as described in text.

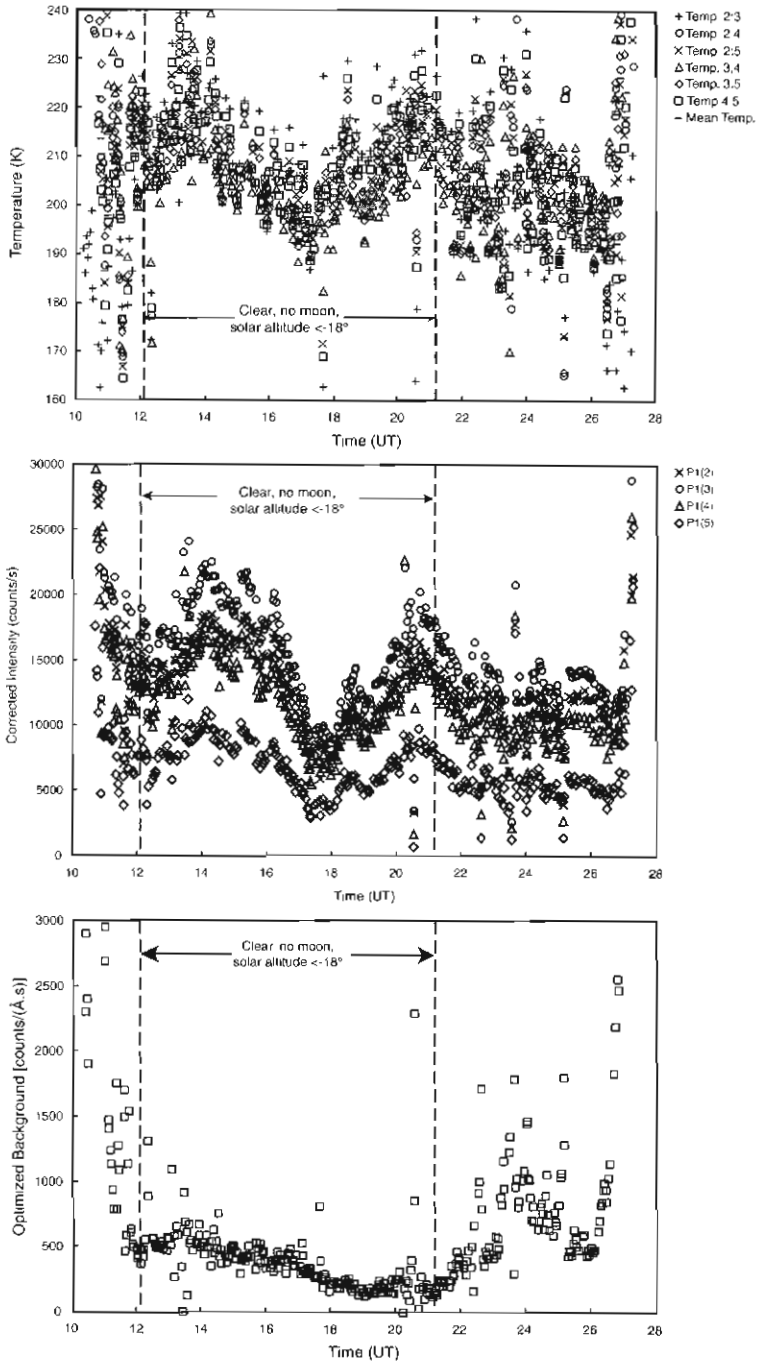


Figure 6. *Davis, 1994: 1000 UT–Day 182 to 0400 UT – Day 183*
(a) Temperatures derived from individual ratios, and their mean.
(b) Intensities of the major P-branch lines.
(c) Background, calculated as described in text.

7.5 How to progress from here

1. Determine the fraction of the background filter counts that best matches the background determined by minimising the standard deviation of temperatures derived from the six possible major line ratios.

2. Determine temperatures and associated temperature-equivalent, counting uncertainties (defining the temperature variation limit) for intervals of no moon, Sun more than 18° below the horizon, and clear skies.

3. Determine the spectral profile of the temperatures measured. Determine the percentage of time when the power in selected bands significantly exceeds the power in the counting noise, and whether any seasonal or diurnal dependence is apparent.

4. Determine the period, magnitude and duration of any temperature fluctuations down to a 15 minute period.

5. Similarly, determine the period, magnitude and duration of intensity fluctuations down to a 15 minute period.

6. Determine the phase relationship between the temperature and intensity fluctuations, for times when the frequencies match.

Reference

Greet, P.A., French, W.J.R., Burns, G.B., Williams, P.F.B., Lowe, R.P. and Finlayson, K. (1998). OH(6-2) spectra and rotational temperature measurements at Davis, Antarctica. *Annales Geophysicae*, 16: 77–89.

8. TIDAL PERIODICITIES IN OBSERVATIONS OF THE OH(6-2) EMISSION FROM EASTERN ANTARCTICA

P.A. Greet [1], D.J. Murphy[1] and P.L. Dyson[2]

[1] Atmospheric and Space Physics
Australian Antarctic Division
Kingston Tasmania 7050
(email: pene.greet@aad.gov.au)

[2] Department of Physics
La Trobe University
Bundoora Victoria 3083

Abstract

High-resolution Fabry-Perot spectrometer observations of the OH(6–2) Q1(1) line, λ834.46 nm, were made at Mawson, 67.6°S 62.9°E, and Davis, 68.6°S 78.0°E, Antarctica, in 1997. The observing program resulted in three sets of continuous OH observations suitable for tidal analysis: July 25–31 when the Mawson FPS observed for seven days, and August 3–7 and 10–14 when the Davis FPS observed for two five day intervals. Lomb-Scargle periodograms show a definitive 12 hour periodicity in the meridional and zonal winds at both stations. The OH emission intensity has a dominant 24 hour periodicity which may be either tidal or photochemical in nature. Superposed epoch analysis is used to determine amplitude and phase of the 12- and 24-hour periodicities. At these latitudes the semi-diurnal tide is dominant with ampltiudes of 10–30 m s^{-1} while the diurnal tide has amplitudes of only 0–5 m s^{-1}. Comparison is made with MF radar measurements. Simultaneous measurements from a radar at Davis compare well with FPS measurements; amplitudes and phases agree and both instruments measure a larger amplitude for the semi-diurnal tide in the last two weeks of July than in the first two weeks of August.

8.1 Introduction

The middle atmosphere is highly variable. Instantaneous measurements of winds, temperature, density, or airglow emission intensity can vary significantly from the mean due to wave or tidal perturbations. The amplitude and phase of waves and tides vary and considerable effort has been made to build tidal models for the mesosphere and lower thermosphere (e.g. Forbes, 1982; Forbes *et al.*, 1997; Yudin *et al.*, 1997). To model the tides properly it is necessary to know the typical tidal variations, the range of variation likely, and ideally the cause of large tidal variations. This requires much data, from a number of techniques, over a period of time. WINDII and HRDI experiments aboard the UARS

satellite have provided data on mesosphere and lower thermosphere tides (e.g. McLandress *et al.*, 1996; Palo *et al.*, 1997). However sampling between 40 and 60 degrees latitude is limited and no sampling occurs above 60 degrees latitude. Tidal observations above 60 degrees are thus of particular interest.

Mesospheric tides are global in nature and specific modes dominate in particular circumstances (Forbes, 1982). At low- and mid-latitudes migrating tidal modes dominate. These modes have either zero or small amplitudes at high-latitudes. Observations at high latitudes indicate the presence of large-scale waves and non-migrating tides (Palo *et al.*, 1998; Sivjee and Walterscheid, 1994; Hernandez *et al.*, 1995, 1996). Mawson and Davis are at latitudes still affected by migrating tides (Forbes, 1982; Phillips and Vincent, 1989; Greet and Dyson, 1999). It is possible that several tidal modes may be present simultaneously.

Tides are an important part of mesospheric dynamics and knowledge of their behaviour is necessary, for example, if effects of long-term climate change are to be detected. Systematic bias due to changes in tidal amplitude or phase, inherent in the data sampling, must be allowed for or considered.

This paper presents Fabry-Perot spectrometer observations of the OH (6–2) Q1(1) line at λ834.5 nm. Observations of this emission commenced at Mawson in 1993 (Greet *et al.*, 1994). A subsequent campaign in 1995 revealed semi-diurnal tidal periodicities in winds and intensities (Greet and Dyson, 1999). Further OH observations from Mawson, 67.6°S 62.9°E, and Davis, 68.6°S 78.0°E, in 1997 also contain tidal periodicities as discussed in the following sections.

8.2 Instruments and data

The Mawson Fabry-Perot spectrometer is described in detail by Jacka (1984). A similar instrument, originally operated at Mt Torrens in the Adelaide hills (Jacka *et al.*, 1980; Cocks *et al.*, 1980), has been adapted for dual-channel, single-etalon operation and installed at Davis. Routine observations at Davis commenced in 1997. OH observing campaigns are co-ordinated between the two stations. Greet *et al.* (1994) describe the technique used to determine the relative intensity of the OH nightglow emission, and mesospheric winds and temperature variations from FPS observations of the (6–2) Q1(1) line at λ834.5 nm.

In 1997 an OH campaign was scheduled for the month of July at Mawson. With the dual-channel operation at Davis, OH observations were scheduled for a longer interval, the month of July and continuing until mid-August. For FPS OH observations to be carried out the weather must be fine and the sky mostly clear. At times observations are made through thin cloud but if cloud thickened or there was any chance of snow, observations ceased. This paper will look in detail at three intervals: July 25–31 when the Mawson FPS observed for seven days, and August 3–7 and 10–14 when the Davis FPS observed for two five-day

intervals. In 1997 weather conditions did not permit the two stations to observe simultaneously for more than two consecutive days.

To extract tidal information from highly variable mesospheric measurements a number of days of data are required. Tides vary in phase and amplitude and analysis techniques assume the phase and amplitude are constant. This is reasonable for periods of about 10 days. Optical observations are limited to dark skies and hence data are available for considerably less than 24 hours a day, at all but the highest latitudes. Even so, tidal variations are obvious in mesospheric observations but care must be taken to consider aliasing effects of optical sampling.

Palo *et al.* (1997) discuss the effects of aliasing on sampled data. Aliasing masks, similar to those in Figure 2 of Palo *et al.* (1997) were generated for the 1997 FPS OH campaigns. The masks are Fourier transforms of step functions to select data samples from a uniformly sampled population. The mask for meridional samples of the Davis August 10–14 campaign is shown in Figure 1. All masks had similar characteristics with substantial power at one cycle per day, and further less-significant power peaks at multiples of one cycle. Palo *et al.* (1997) note that this can cause leakage of power from 24 hours into the mean, from the mean and 12 hours into the diurnal tide and from 24 and 8 hours into the semi-diurnal tide.

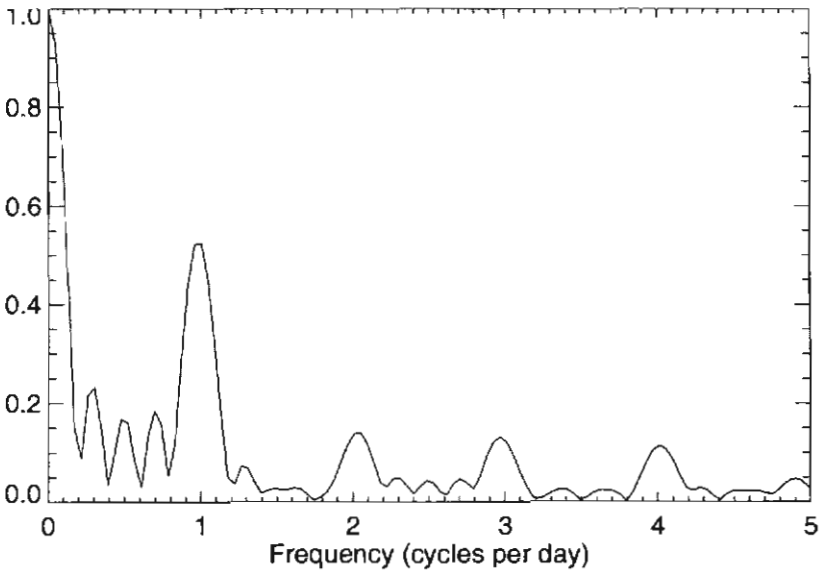


Figure 1. *Aliasing spectrum from meridional samples during the Davis campaign 10–14 August 1997. Significant power occurs at one cycle per day, and some power at 2, 3, and 4 cycles per day.*

In 1995, an OH observing campaign at Mawson obtained 14 days data in a 15-day interval. This has been analysed by Greet and Dyson (1999). In this data it was possible, using Lomb-Scargle periodograms, to identify the semi-diurnal tide in the meridional wind data. A Lomb-Scargle periodogram of raw data was performed, followed by a second periodogram of the raw data minus a 12- or 24-hour fit. Simulations showed that if the wrong periodicity were subtracted from the data, power at eight hours would remain in the periodogram. This type of analysis for the 1997 campaigns will be presented in the next section.

No significant periodicities were present in temperature measurements from the 1995 campaign and none were found in any of the three 1997 campaigns. This does not imply that no temperature periodicities are present, merely that any temperature variation is less than ~ 11 K, approximately half the mean uncertainty in the temperature measurements (Greet and Dyson, 1999). Temperature measurements will not be included in the following discussion.

8.3 Lomb-Scargle analysis

Figure 2 shows Lomb-Scargle periodograms of the OH emission intensity and wind, for meridional and zonal measurements from the Mawson campaign, July 25–31 1997. Each of the seven days in this interval had 14–15 hours of data. Periodograms of the raw data, column (a) of Figure 2, show that 11.9 ±0.5 hour is the most significant periodicity in the meridional and zonal wind, although it is not statistically significant in the meridional wind. There is no significant power at 24 hours in either the meridional or zonal wind. In the zonal wind, subtraction of a fitted 12-hour periodicity, column (c), leaves no significant periodicities and subtraction of a fitted 24-hour periodicity, column (d), leaves some power at 8-hour, albeit only at a 50% significance level. This is evidence of a 12-hour periodicity in the wind.

In both the meridional and zonal emission intensity the most significant periodicity is near 24 hours. There is some power at 12 hours in the meridional emission intensity but it is not significant, and in the zonal intensity the power at 12 hours is near the noise level. Subtracting a fitted 12-hour periodicity leaves power at 8 hours, again not reaching the 95% significance level. Subtracting a 24-hour periodicity leaves power only at longer periodicities. This is evidence of a 24-hour periodicity in the intensity.

OH emission intensity may show periodicities from photochemical as well as tidal sources.

Intensity measurements during the Davis campaign from 3–7 August 1997, are dominated by tropospheric cloud which generated a 2.5 day periodicity. There was significant power at 21 ±1 hours in the meridional and zonal emission intensity but no significant power at or near 12 hours. In the meridional (zonal) wind, 11 ±1 (11.6 ±0.6) hours was the most significant periodicity; 23 ±2 hours was also significant in the zonal

wind. Subtraction of fitted 12- or 24-hour periodicities did not distinguish between them. Periodicities from this interval are slightly less, although not always significantly less, than theoretical tidal periodicities of 12 and 24 hours.

The Davis campaign from 10–14 August 1997 also had significant power at 22 ±1 hours in the zonal and meridional emission intensity. Power at 12 hours was significant at the 50% level in the meridional emission intensity, and almost so in the zonal emission intensity. In the meridional wind 12 hours was the most significant periodicity and 26 ±3 hours was also significant. In the zonal wind 12 hours was

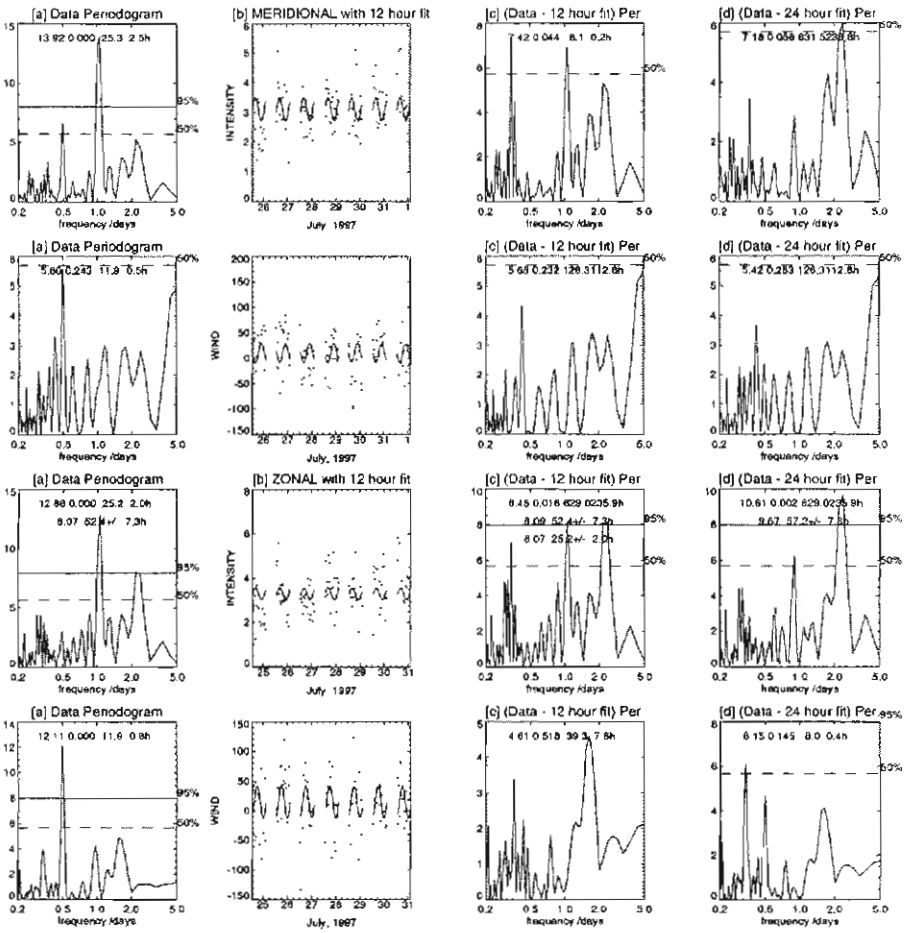


Figure 2. *Lomb-Scargle periodograms of meridional emission intensity and wind, and zonal emission intensity and wind for the Mawson campaign, July 25–31 1997. (a) Periodogram of data, (b) data with 12 hour fit, (c) periodogram of data with 12 hour fit subtracted, and (d) periodogram of data with 24 hour fit subtracted. After subtraction of a fitted component, residual power at 8 hours indicates an inappropriate fit to the data.*

the most significant periodicity, but not quite reaching the 95% level of significance. Subtracting 12 and 24 hour fitted periodicities did not provide evidence for one periodicity over another. This is possibly due to a contribution from both 12 and 24 hours, with 12 hours being dominant in the wind and 24 hours being dominant in intensity variations.

Power at a given periodicity does not necessarily imply the presence of a tidal oscillation. If the phases of meridional and zonal periodicities are in agreement with that predicted by models then we have more confidence in a tidal interpretation. Attempts to obtain amplitude and phase information from simultaneous fitting of 12- and 24-hour periodicities will now be discussed.

8.4 Superposed epoch analysis

Previous studies at Mawson (Vincent, 1994) suggest that, at this time of year, the semi-diurnal tide is dominant and the contribution of the diurnal tide is small.

Interannual variability shows that at times the diurnal tidal contribution may be significant. Lomb-Scargle analysis suggests that, for the intensity at least, diurnal variations are significant or dominant. Greet and Dyson (1999) only fitted either 12- or 24-hour components to the data. If both 12- and 24-hour components are present then ideally they should be fitted simultaneously.

Fitting tidal components to the data also suffers from the effects of aliasing. The Lomb-Scargle analysis used a simple IDL curve-fitting routine, CURVEFIT, to subtract the mean plus 12- or 24-hour components. This will be called a two component fit. This was extended and a more robust curve fitting routine, SVDFIT, was also used to fit a mean, 12- and 24-hour (three) components. Both of these routines gave problems, with power from the mean leaking into the 24-hour component being the most common fault. This is obvious comparing a two component, three parameter, fit and the three component, five parameter, fit. Anomalously low, or high, mean value and high diurnal value are returned by the latter.

Large point-to-point variability hinders fitting tidal components. High frequency variability is common to most mesospheric measurements. A superposed epoch analysis eliminates, or at least minimises, high-frequency variability. Figure 3 shows the analysis for the combined Davis meridional and zonal winds. In the combined Davis data set, 10 of 12 days in the interval 3–14 August are sampled. Column (a) contains periodograms, (b) superposed epoch analysis fitting mean and combined 12- and 24-hour, and (c) as for (b) but only fitting mean and 12-hour components. Because of the complexity of the intensity spectra, the dominance of 24-hour rather than 12-hour components, with photochemical and tidal effects present, the following analysis deals only with wind measurements.

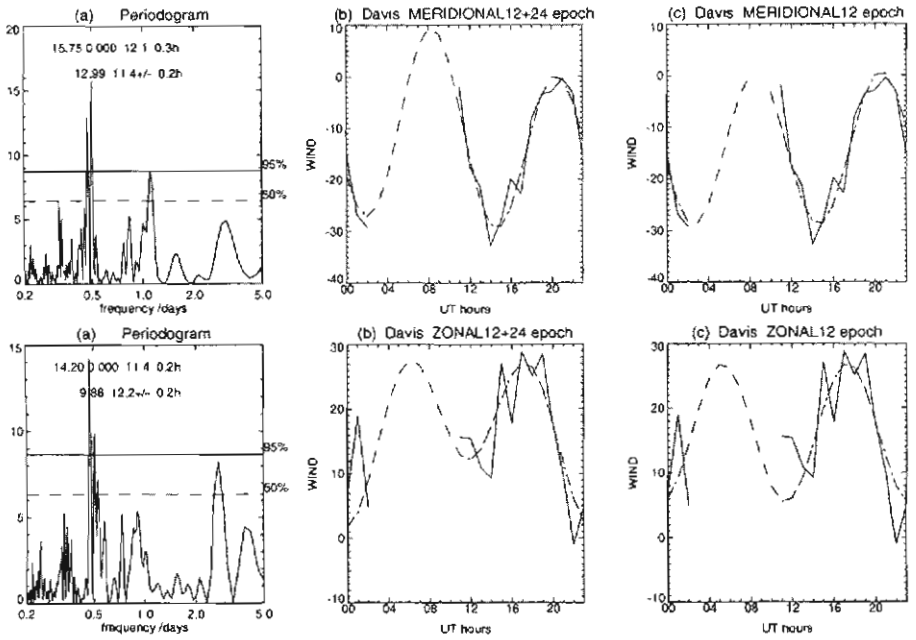


Figure 3. *Superposed epoch analysis of the combined Davis wind meridional (top) and zonal (bottom) data sets, 10 days of data in the 12 day period 3–14 August 1997. (a) Periodograms, (b) superposed epoch data, and 5 component mean, 12- and 24-hour fit (dashed curve) (c) as for (b) but three component mean and 12 hour fit.*

The periodograms show two significant frequencies, 12.1 and 11.4 hours, in both the meridional and zonal wind. The aliasing mask for the zonal data from the combined Davis campaigns, in Figure 4, shows high frequency structure which would produce this splitting of the power at 12 hours. The power at 2.12 cycles per day is larger in the zonal than meridional, 0.65 as compared to 0.45, which would explain the difference in the Lomb values for the two data sets.

The fitted diurnal component is small. There is reasonable agreement between the fitted mean and 12-hour components from both the two and three component fits for both the meridional and zonal data. Table 1 gives the fitted parameters for Davis and Mawson data. Fitted values are grey if a poor fit was noted on visual assessment. Other discrepancies are obvious when comparing the two fits, in particular, the amplitude of the 12- and 24-hour oscillations in the Davis 2 meridional data set, and Mawson zonal data set. In both cases aliasing is skewing the results and they have also been colour-coded grey.

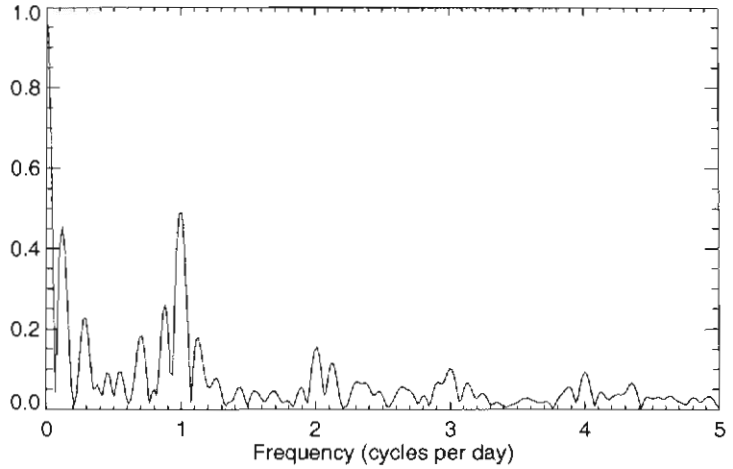


***Figure 4**. Aliasing spectrum from zonal samples during the combined Davis campaign 3–14 August 1997. The small scale structure producing peaks at both 2.0 and 2.12 cycles per day splits the power at 12 hours in the data into two periodicities, as seen in Figure 3. Note differences from Figure 1.*

***Table 1**. Fitted amplitudes, in m s^{-1}, and phases, in hour UT, for ZONAL wind, various campaigns.*

		n_{days}	A_0	A_{12}	Φ_{12}	A_{24}	Φ_{24}	A_0	A_{12}	Φ_{12}
Davis	3–14 Aug	10 in 12	17	10	6	5	12	16	11	5
Davis 1	3–7 Aug	5 in 5	17	5	3	9	13	17	8	4
Davis 2	10–14 Aug	5 in 5	14	14	6	2	22	15	15	6
Mawson	25–31 Jul	7 in 7	0	25	5	31	20	16	27	6

***Table 2**. As for Table 1, but for MERIDIONAL wind.*

		n_{days}	A_0	A_{12}	Φ_{12}	A_{24}	Φ_{24}	A_0	A_{12}	Φ_{12}
Davis	3–14 Aug	10 in 12	-12	17	8	5	7	-14	15	9
Davis 1	3–7 Aug	5 in 5	-14	17	8	8	8	-18	13	8
Davis 2	10–14 Aug	5 in 5	-8	40	9	19	11	-11	50	9
Mawson	25–31 Jul	7 in 7	15	22	10	11	8	11	27	10

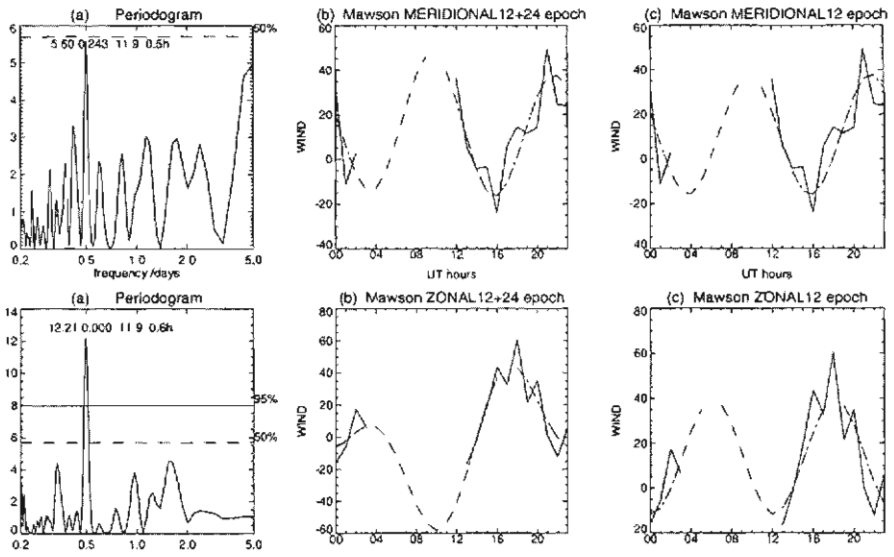


Figure 5. *Superposed epoch analysis of the Mawson wind meridional (top) and zonal (bottom) data sets, from 25–31 July 1997. (a) Periodograms, (b) superposed epoch data, and 5 component mean, 12- and 24-hour fit (dashed curve) (c) as for (b) but three component mean and 12 hour fit.*

Allowing for the diurnal component by fitting five parameters is of limited success. The amplitude and phase of the 12-hour component is in reasonable agreement between the two fits. Figure 5 shows superposed epoch analysis for the Mawson campaign. Expanding the data set to include partial days and isolated nearby days did not significantly alter the result.

8.5 Discussion

8.5.1 High-latitude semi-diurnal tide

The fitted 12-hour component is significantly larger than the 24-hour component. Figure 3 and Tables 1 and 2, show that for the fitted 12-hour component the zonal wind phase leads the meridional by approximately $\pi/2$, 3 hours, as expected for the migrating tide in the southern hemisphere.

'During April, May, July, and August the phases at Scott Base, Molodezhnaya, and Mawson are within about 1 hour of 0100 LT; a constant phase with respect to longitude is consistent with a semi-diurnal oscillation migrating with the apparent motion of the sun.' Portnyagin *et al.* (1998) note when comparing the 12-hour oscillation, from radar observations, at South Pole, Scott Base, Molodezhnaya and Mawson. The phase referred to here is that of the meridional component.

From the two component fit in Figure 3 the Davis meridional 12-hour component maximises near 0800 UT, which is 0100 LT to within 15 minutes. From similar plots for Mawson, Figure 5, the meridional 12-hour component maximises near 1000 UT, which is 0200 LT to within 15 minutes. These values for Mawson and Davis are not from simultaneous data sets. A superposed epoch analysis was performed on data obtained simultaneously at Davis, the Davis FPS observed for 8 of the 14 days from July 19–August 1, but there were insufficient data to well fit amplitude and phase parameters.

To better compare phases between Mawson and Davis we need to split the data sets into individual directions as Mawson east and Davis west viewing volumes are adjacent to each other compared to the distances between east and west viewing volumes for one individual station. This was done but with marginal success because of poor signal-to-noise. Values obtained from the superposed epoch fit for all Davis and Mawson data in individual directions show that the amplitude of the 12-hour component maximises in the Davis east viewing volume, approximately two hours prior to the Mawson west viewing volume. These two volumes are separated by 30 degrees of longitude. Rigorous uncertainty calculations have not been included and considering the limited number of points this result should be considered qualitatively rather than quantitatively. That the Davis east volume phase is in advance of the Mawson phase is further evidence for interpretation of the 12-hour oscillation as the migrating semi-diurnal tide.

This discussion of the semi-diurnal tide has been restricted to oscillations in the wind. It is hoped that by comparison with models we can extend our analysis to include OH emission intensity. At present an approximately 24-hour periodicity dominates in the OH emission intensity. There is some power at 12 hours but aliasing and the presence of both photochemical and tidal variations make interpretation of the OH emission intensity data beyond the scope of this work.

8.5.2 Comparison with Mawson radar winds

An MF radar was operational at Mawson from 1981–93 and data are readily available for the years 1984–90 (Vincent, 1994). The radar was upgraded and moved to Davis in 1994 and was operational during 1997. The radar samples every two kilometres with a four kilometre range resolution.

The OH emission originates from a layer of approximately 7 km half-width at a height of 87 km (Baker and Stair, 1988). The layer shape varies significantly with large tidal and wave perturbations in the region. FPS observations integrate the OH emission over the entire layer. By ignoring possible complexities in the OH layer shape a simple comparison can be made between the FPS OH observations and a single radar height near the peak of the OH layer. We have chosen a radar height of 86 km for such a comparison.

Spatial variations also potentially affect measurements from two instruments. The radar beam was 40 degree half-width half-maximum. This corresponds to a field-of-view of approximately 150 km at an altitude of 87 km. The radar beam pointed to the zenith. The FPS semi-angle field-of-view was 13 mrad. This corresponds, at a 75 degree zenith angle, to a column through the OH layer approximately 40 km in diameter. The meridional and zonal FPS viewing volumes are separated by 650 km. Thus the two instruments spatially average in quite different ways.

Meridional and zonal winds from the FPS's, the MF radar and a model are compared in Tables 3 and 4. These tables are expansions of Tables 1 and 2 to include: simultaneous MF radar data from Davis for the specific intervals of the FPS campaigns; 12 day radar averages from Mawson for the years 1984–90 (Vincent, 1994); and GSWM-98 model values for a height of 86 km and latitudes 66°S and 69°S (Hagan *et al.*, 1999; http://www.hao.ucar.edu/public/research/tiso/gswm/gswm.html). The values in square brackets next to the Mawson radar averages are the standard deviations indicative of typical variations in the mean.

The Mawson FPS mean and semi-diurnal tidal amplitudes are significantly larger than the average Mawson MF radar values. The Mawson radar values are averaged over seven years and, although an indication of variability is available, the value obtained on any given year may differ substantially. Considering the spatial differences in the two sampling techniques it is probably still useful to also compare the Mawson FPS amplitudes with those obtained simultaneously by the Davis radar. Theoretically the phase of the tide at Mawson should lag that at Davis by one hour. We will first consider the diurnal and then the semi-diurnal tides.

The fitted 24-hour component in both the FPS and radar winds is small, of the order of 5 m s^{-1} or less. It is difficult for the FPS to well determine the amplitude and phase of such a small contribution considering the inherent optical sampling difficulties and general mesospheric variability. Meridional and zonal amplitude and zonal phase values obtained from the three component fit are in reasonable agreement with those of MF radar and the GSWM model. Diurnal variations in OH emission intensity are larger and hence probably photochemical in origin.

The amplitude of the semi-diurnal tide is also more similar in the FPS and simultaneous Davis radar than the averaged Mawson radar amplitudes. The FPS 12-hour amplitudes are slightly larger than the Davis radar amplitudes possibly because of the smaller FPS sampling volumes. Differences between the radar and FPS semi-diurnal amplitudes are also associated with differences in the fitted mean winds.

Table 3. *Fitted amplitudes, m s^{-1}, and phases, in hour UT, for ZONAL wind, various campaigns.*

		n_{days}	A_0	A_{12}	Φ_{12}	A_{24}	Φ_{24}	A_0	A_{12}	Φ_{12}
Mawson	25–31 Jul	7 in 7	0	25	5	31	20	16	27	6
Davis	3–14 Aug	10 in 12	17	10	6	5	12	16	11	5
Davis 1	3–7 Aug	5 in 5	17	5	3	9	13	17	8	4
Davis 2	10–14 Aug	5 in 5	14	14	6	2	22	15	15	6
Davis	25-31 Jul	Radar	6	18	6	1	7			
	3–14 Aug	Radar	16	10	6	1	12			
	3–7 Aug	Radar	29	11	6	4	5			
	10–14 Aug	Radar	10	14	6	5	16			
Mawson	204, 84–90	Radar	10[6]	9[3]	7[5]	7[7]	9[6]			
	216, 84–90	Radar	8[8]	9[5]	5[6]	5[3]	11[6]			
	228, 84–90	Radar	11[8]	9[5]	9[5]	4[3]	12[7]			
GSWM	July solstice	Lat 66°S		22.7	12.1	4.1	18.3			
		Lat 69°S		20.3	12.2	6.6	17.7			

Table 4. *As for Table 3 but for MERIDIONAL wind, various campaigns.*

		n_{days}	A_0	A_{12}	Φ_{12}	A_{24}	Φ_{24}	A_0	A_{12}	Φ_{12}
Mawson	25–31 Jul	7 in 7	15	22	10	11	8	11	27	10
Davis	3–14 Aug	10 in 12	-12	17	8	5	7	-14	15	9
Davis 1	3–7 Aug	5 in 5	-14	17	8	8	8	-18	13	8
Davis 2	10–14 Aug	5 in 5	-8	40	9	19	11	-11	50	9
Davis	25-31 Jul	Radar	-6	19	8	5	4			
	3–14 Aug	Radar	-10	12	8	3	3			
	3–7 Aug	Radar	-12	13	8	2	4			
	10–14 Aug	Radar	-8	17	8	3	1			
Mawson	204, 84–90	Radar	3[4]	10[5]	5 [4]	6[4]	9[9]			
	216, 84–90	Radar	3[4]	11[6]	4[4]	5[4]	9[6]			
	228, 84–90	Radar	-1[5]	12[7]	3[2]	7[7]	10[9]			
GSWM	July solstice	Lat 66°S		23.4	3.1	5.7	0.1			
		Lat 69°S		21.8	3.1	5.5	0.2			

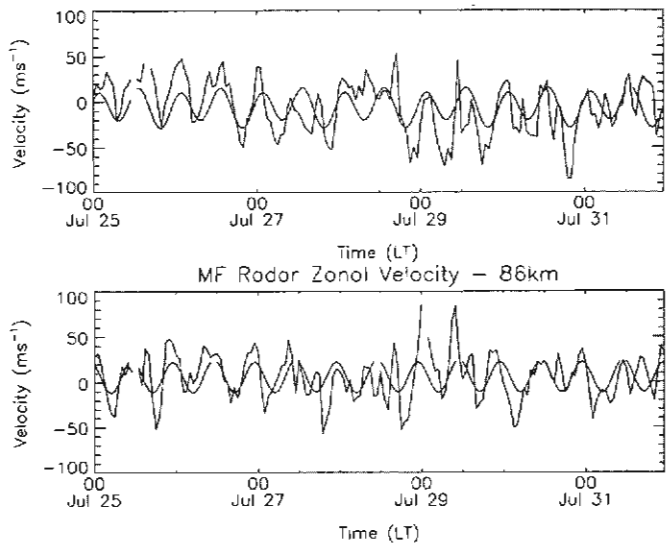


Figure 6. *Davis MF radar meridional and zonal winds obtained during 25–31 July when the Mawson FPS was observing. The smooth overlaid curve is the combined mean, 12- and 24-hour fitted components*

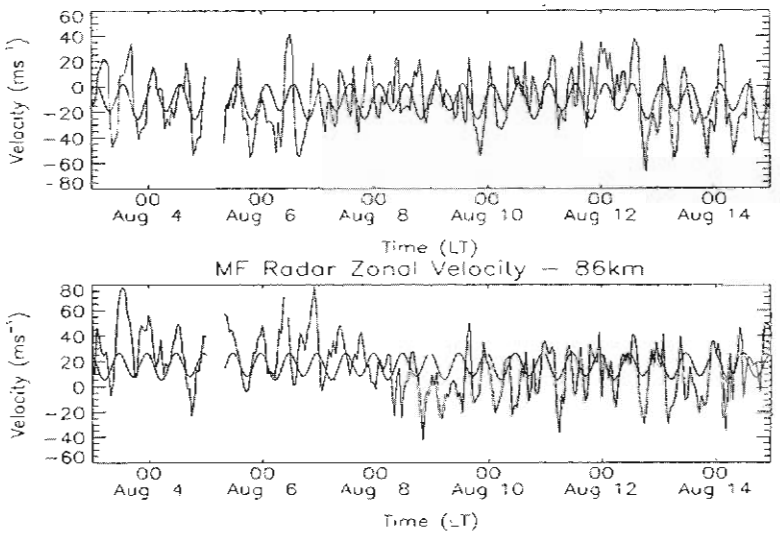


Figure 7. *Davis MF radar meridional and zonal winds obtained during 3–24 August 1997. This coincides with the two Davis FPS campaigns. The smooth overlaid curve is the combined mean, 12- and 24-hour fitted components.*

There is good agreement between the general trends in the FPS and radar mean and 12-hour components. Both the simultaneous Davis radar and the FPS measurements indicate that the amplitude of the 12-hour component is larger in the last two weeks of July than in August. The mean zonal and meridional winds vary in a similar fashion between the Davis radar and FPS observations, although the Mawson FPS means are somewhat different.

The well fitted zonal FPS mean and 12-hour components are in good agreement, both between the two and three component FPS fits and the FPS and radar fits. For the meridional winds, the two and three component FPS fits vary more and the agreement between the FPS and simultaneous Davis radar values is not as good as for the zonal winds. This may be due to variability in the tide and/or the presence of other perturbations in the wind field which makes the tides difficult to isolate. Figures 6 and 7 show the Davis radar measured meridional and zonal winds for the Mawson and combined Davis campaigns. The continuous radar sampling allows easier visual identification of tidal variations. The smooth overlaid curve is the combined, mean, 12- and 24-hour fit.

8.5.3 Comparison with other southern polar wind studies

Hernandez *et al.* (1995) looked at periodicities in FPS OH observations at South Pole and MF radar at Scott Base identifying 10-hour and 2-day periodicities in Scott Base and South Pole data, as well as 7-, 12- and 24-hours at Scott Base. South Pole data were used to identify the oscillations as zonal wavenumber one waves. Portnyagin *et al.* (1998) have studied in detail periodicities in meteor wind data from South Pole over a year. They note the dominance of 12- and 24-hour periodicities in summer and from April to October the dominance of oscillations with periods less than 12 hours and periods from 2–10 days. Sivjee and Walterscheid (1994) identified temperature oscillations of six hours in OH emission brightness and rotational temperature measurements at South Pole. We note that in FPS wind measurements from Mawson and Davis there are no significant periodicities other than 12 hours.

8.6 Conclusions

FPS observations of the OH emission have been made at Mawson and Davis during 1997. Lomb-Scargle periodograms have been used to identify a 12-hour periodicity in the wind. A superposed epoch analysis has been used to fit mean, 12- and 24-hour components to the data. As the zonal phase of the 12-hour component leads the meridional phase by three hours, and the phase at Davis leads that at Mawson, and in both cases maximises at approximately 0100 LT, we identify this oscillation as the migrating semi-diurnal tide. Fitted amplitudes of the diurnal tide are of the order of or less than 5 m s^{-1}. The amplitude of the semi-diurnal tide varies between 10 and 30 m s^{-1} over the campaigns.

Comparisons have been made between the FPS and MF radar measurements. Averaged radar tides from seven years of data at Mawson imply smaller values than were seen by the FPS. Simultaneous data

obtained with the MF radar at Davis were in good agreement with the FPS measurements. Both instruments measured larger, 20–30 m s^{-1}, semi-diurnal tides in the last two weeks of July than in August when the amplitudes were in the range 10–15 m s^{-1}. Differences between the two instruments may be due to the different spatial averaging or different fitting routines. The times when differences between the two techniques were largest coincided with the times when the tides were less obvious or more variable.

The only frequencies positively identified in this study are those at 12- and 24-hours. There is no indication of the presence of 10-hour oscillations in the three 1997 campaigns.

Further OH observing campaigns have been carried out and it is hoped that the FPS and radar comparison can be extended to larger data sets and times when the Mawson and Davis FPS's obtained simultaneous data.

8.7 Acknowlegements

We acknowledge support by the Australian National Antarctic Research Expeditions, the Australian Antarctic Science Advisory Committee, and the Australian Research Council. Mr Chris Boucher collected the 1997 Mawson FPS data and Mr Mike Manion assisted with the Davis FPS and ran the Davis radar in 1997. General support from other expeditioners through the year was appreciated. Dr R.A. Vincent, University of Adelaide, supplied the Mawson radar tidal data. Thanks to M. Hagan and associates for free, easy web access to the GSWM-98 results at the site: http://www.hao.ucar.edu/public/research/tiso/gswm/gswm.html.

References

Baker, D.J. and Stair, A.T. (1988). Rocket measurements of the altitude distributions of the hydroxyl airglow. *Physica Scripta,* 37: 611–622.

Cocks, T.D., Creighton, D.F. and Jacka, F. (1980). Application of a dual Fabry-Perot spectrometer for daytime airglow studies. *Journal of Atmospheric and Terrestrial Physics,* 42: 499–511.

Forbes, J.M. (1982). Atmospheric tides 2. The solar and lunar semidiurnal components. *Journal of Geophysical Research,* 87: 5241–5252.

Forbes, J.M., Hagan, M.E., Zhang, X. and Hamilton, K. (1997). Upper atmosphere tidal oscillations due to latent heat release in the tropical troposphere. *Annales Geophysicae,* 15: 1165–1175.

Greet, P.A. and Dyson, P.L. (1999). Tidal periodicities in observations of the OH(6–2) emission from Mawson, Antarctica. *Advances in Space Research,* 24: 579–582.

Greet, P.A., Innis J.L. and Dyson P.L. (1994). High resolution Fabry-Perot observations of mesospheric OH(6-2) emissions. *Geophysical Research Letters,* 21: 1153–1156.

Hagan, M.E., Burrage, M.D., Forbes, J.M., Hackney, J., Randel, W.J. and Zhang, X. (1999). GSWM-98: Results for migrating solar tides. *Journal of Geophysical Research,* 104: 6813–6828.

Hernandez, G., Smith, R.W. and Fraser, G.J. (1995). Antarctic high-latitude mesospheric dynamics. *Advances in Space Research,* 16(5): 71–80.

Hernandez, G., Forbes, J.M., Smith, R.W., Portnyagin, Y., Booth, J.F. and Makarov, N. (1996). Simultaneous mesospheric wind measurements near South Pole by optical and meteor radar methods. *Geophysical Research Letters,* 23: 1079–1082.

Jacka, F. (1984). Application of Fabry-Perot spectrometers for measurement of upper atmosphere temperatures and winds. In: Vincent, R. (Ed.). *Middle Atmosphere Program,* 13: 19–40.

Jacka, F., Bower, A.R.D., Creighton, D.F. and Wilksch, P.A. (1980). A large-aperture high-resolution Fabry-Perot spectrometer for airglow studies. *Journal of Physics E: Scientific Instrumentation,* 13: 562–568.

McLandress, C., Shepherd, G.G. and Solheim, B.H. (1996). Satellite observations of thermospheric tides: results from the WIND imaging interferometer on UARS. *Journal of Geophysical Research,* 101: 4093–4114.

Palo, S.E., Hagan, M.E., Meek, C.E., Vincent, R.A., Burrage, M.D., McLandress, C., Franke, S.J., Ward, W.E., Clark, R.R., Hoffmann, P., Johnson, R., Kürschner, D., Manson, A.H., Murphy, D., Nakamura, T., Portnyagin, Yu., I., Salah, J.E., Schminder, R., Singer, W., Tsuda, T., Virdi, T.S. and Zhou, Q. (1997). An intercomparison between GSWM, UARS, and ground based radar observations: a case-study in January 1993. *Annales Geophysicae,* 15: 1123–1141.

Palo, S.E., Portnyagin, Y.I., Forbes, J.M., Makarov, N.A. and Merzlyakov, E.G. (1998). Transient eastward-propagating long-period waves observed over South Pole. *Annales Geophysicae,* 16: 1486–1500.

Portnyagin, Y.I., Forbes, J.M., Makarov, N.A., Merzlyakov, E.G. and Palo, S.E. (1998). The summertime 12-h wind oscillation with zonal wavenumber s = 1 in the lower thermosphere over the South Pole. *Annales Geophysicae,* 16: 828–837.

Phillips, A. and Vincent, R.A. (1989). Radar observations of prevailing winds and waves in the southern hemisphere mesosphere and lower thermosphere. *PAGEOPH,* 130: 303–318.

Sivjee, G.G. and Walterscheid, R.L. (1994). Six-hour zonally symmetric tidal oscillations of the winter mesopause over the South Pole Station. *Planetary and Space Science*, 42: 447–453.

Vincent, R.A. (1994). Gravity-wave motions in the mesosphere and lower thermosphere observed at Mawson, Antarctica. *Journal of Atmospheric and Terrestrial Physics*, 56: 593–602.

Yudin, V.A., Khattatov, B.V., Geller, M.A., Ortland, D.A., McLandress, C. and Shepherd, G.G. (1997). Thermal tides and studies to tune the mechanistic tidal model using UARS observations. *Annales Geophysicae*, 15: 1205–1220.

9. MULTIPLE PARAMETER SIMULATION OF THE FABRY-PEROT INTERFEROMETER

T. Davies and P.L. Dyson

Physics Department
Faculty of Science, Technology and Engineering
La Trobe University
Bundoora Victoria 3083 Australia
(email: t.davies@latrobe.edu.au)

Abstract

Two simulation programmes have been written as aids to the design of Fabry-Perot interferometers. The use of the Modulation Transfer Function (MTF) vis-a-vis the Aperture Function is discussed and an example given. Other features of the simulations are described briefly.

Natural broadening and collisional broadening of a spectral line gives rise to a Lorentzian intensity distribution and Doppler broadening gives a Gaussian intensity distribution (Gill, 1965). The shape of the airglow spectral line is given by the convolution of a Lorentzian and Gaussian function which is called a Voigt function (Rees, 1989). The Voigt profile is Gaussian from the peak out to about three full widths at half height, from the peak (Thorne, 1974), so for the purpose of this discussion the airglow spectral line may be considered to have a Gaussian shape.

The shape of the recorded spectral line is further broadened by the interferometer instrument function. The instrument function is composed of a function representing the line shape due to the multiple reflections in the etalon and subsequent interference.

$$y(x) = y_{max}\left[1 + \left(\frac{2FRS}{\pi FWHH}\right)^2 sin^2(Sx - Sp)\right]^{-1}$$

An Airy function is normally used to describe the intensity of the interference fringes (Lothian, 1975; Hecht an Zajac, 1974) but a combination of Lorenzian functions can be used (Vaughan, 1989). The spectrum can be modified by a Gaussian function representing imperfections in the etalon and aberrations in the rest of the optics. If Lorenzian functions are used to describe the action of the etalon, then the recorded spectrum can be modelled with a Voigt function (Vaughan, 1989). However, the usual procedure is to model the recorded spectrum using a convolution of Airy and Gaussian functions.

In a non-imaging (scanning) interferometer there is a broadening of the recorded spectral line due to the effect of the sampling aperture or field-

stop. An aperture function is normally expressed in terms of the fraction of an order of interference, δm, admitted by the aperture or field-stop (Hernandez, 1986). Such a function is given by Wilksch (1975) as:

$$\delta m = \frac{2nt\delta \cos\theta}{\lambda}$$

θ is the angle between an accepted ray and the normal to the reflecting surfaces of the etalon plates, and δθ is the range of angles accepted by the aperture.

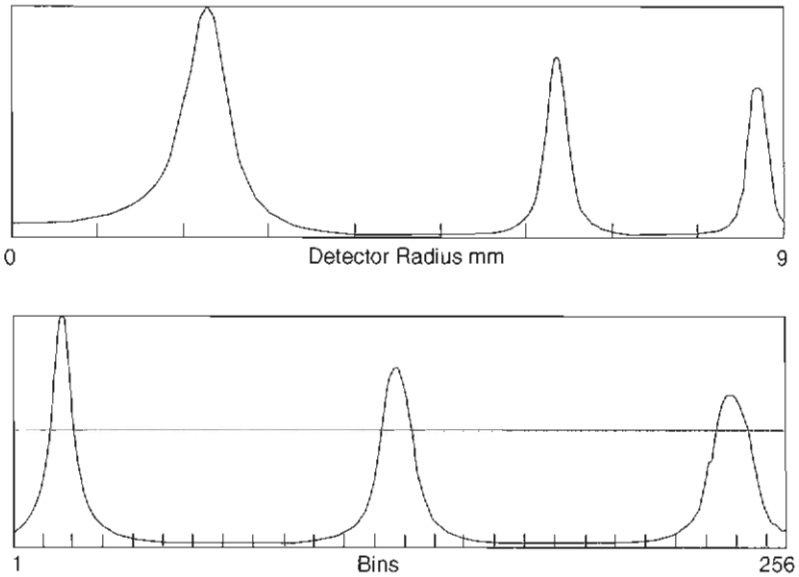


Figure 1. *The upper plot shows recorded intensity as a function of detector radius for three interference fringes. The lower plot shows the spectrum that would be recorded as successive fringes are lowered and widened by the MTF. Small glitches are the result of binning and mapping.*

For a scanning instrument δθ is fixed and the fractional change in t is very small, so the aperture function is essentially fixed. For an imaging interferometer δθ is defined by each pixel. For a constant δm, δθ will change with the distance between each bin or channel of the spectrum. The cone of rays admitted by the field-stop is much wider than that accepted by a single pixel, so a better way to describe broadening by the detector is to use the MTF of the imaging detector. The MTF of an Imaging Photon Detector has been modelled numerically and the broadening of spectral peaks is shown in Figure 1. The bin number is proportional to the square of the radius. The peaks would all have the same width in bins if the MTF of the detector and optics was always unity.

In addition to the MTF the effect of temperature and pressure changes of air (nitrogen) in the etalon chamber have been modelled using the equation (Kaye and Laby, 1972):

$$n_{TP} = (n-1) \times \frac{P\left(1 + P(61.3 - T) \times 10^{-10}\right)}{96095.4(1 + 0.003661T)}$$

Other design parameters which can be varied in simulations are: focal length, etalon spacing, air pressure, air temperature and instrument finesse. In addition it is possible to simulate changes in wavelength and temperature of airglow emissions.

Definitions of symbols:

FSR = free spectral range
FWHH = full width at half height
m = order of interference
λ = wavelength of light
n = refractive index of medium (air) between etalon plates
P = air pressure
p = position of first spectral peak on horizontal axis of spectrum expressed in bins or channels
S = scaling factor
T = air temperature
t = spacing between etalon plates
x = position on horizontal axis of spectrum expressed in bins or channels

9.1 Acknowlegement

This work has been supported by an Australian Research Council grant at La Trobe University.

References

Gill, T.P. (1965). *The Doppler Effect*. Logos Press, London.

Hecht, E. and Zajac, A. (1974). *Optics*. Addison-Wesley.

Hernandez, G. (1986). *Fabry-Perot Interferometers*. Cambridge Studies in Modern Optics.

Kaye, G.W.C. and Laby, T.H. (1972). *Tables of physical and chemical constants*. Longman, London & New York.

Lothian, G.F. (1975). *Optics and its uses*. Van Nostrand Reinhold Company, London.

Rees, M.A. (1989). *Physics and chemistry of the upper atmosphere*. Cambridge University Press, Cambridge.

Thorne, A.P. (1974). *Spectrophysics*. Chapman and Hall, London.

Vaughan, J.M. (1989). *The Fabry-Perot interferometer*. Adam Hilger, Bristol.

Wilksch, P.A. (1975). *Measurement of thermospheric temperatures and winds using a Fabry-Perot Spectrometer*. PhD Thesis, University of Adelaide.

10. MODELLING AND MAPPING SPORADIC E USING BACKSCATTER RADAR

R.J. Norman[1], P.L. Dyson[1] and J.A. Bennett[2]

[1] Department of Physics
La Trobe University
Bundoora Victoria 3083 Australia
(email: r.norman@latrobe.edu.au)

[2] Department of Electrical and Computer Science Engineering
Monash University
Clayton Victoria 3168 Australia

Abstract

Sporadic E, abbreviated E_s, is a very thin ionised layer and as its name suggests, it occurs sporadically at E region heights. To a radio wave E_s often behaves as a smooth HF reflector. However, because of its complex structure and large spatial and temporal variability E_s is a major problem, especially, for HF propagation predictions where the main emphasis is on the maximum electron density of the layer and an accurate description of the ionosphere. In this study it will be shown that many features of oblique propagation via E_s can be both successfully modelled and mapped.

10.1 Introduction

The aim of this study is to develop techniques to both model E_s echo traces on, and map E_s from, sporadic E traces on a Backscatter Ionogram (BSI). Thus, these modelling techniques can play an important role in the frequency management system of over-the-horizon radar facilities.

Preliminary work by Houminer *et al.* (1996) proposed that sporadic E clouds or patches could be mapped using Backscatter Sounders (BSS). Their technique has been developed further in this study (see also Norman *et al.*, 1998) such that synthesised backscatter ionograms in conjunction with an ionospheric model containing a sporadic E layer are used to map and model different types of E_s.

The Quasi-Parabolic Segment (QPS) model (Dyson and Bennett, 1988) is an analytic ionospheric model and is commonly used to represent the Earth's ionospheric profile. The QPS model, consists of 5 QPS's, and produces integrable terms when solving for the ray parameters. The five QPS's represent the three ionospheric layers namely the E, F_1 and F_2 layers. The required inputs of the QPS model are the peak plasma

frequencies, the peak heights and the semi-thickness of each of the three ionospheric layers. In this study a further layer is added to the QPS model representing an E_s layer. This new analytic ionospheric model is referred throughout this study as the 4 Layer or 4L model. The only added inputs that this model requires are the peak plasma frequency f_oE_s and the height of the peak plasma frequency of the E_s layer, h_mE_s.

The method of mapping E_s in this paper is very simple and makes use of the thickness, and maximum operating frequency, of the E_s echo traces on the BSI and assumes that the E_s behaves as a spherical mirror. The equations determining the distance of the E_s cloud from the transmitter/receiver location as well as the peak plasma frequency and the extent of the E_s clouds are presented. The effects of patches of E_s and some examples of partial reflection from E_s are also presented.

Two example backscatter ionograms from the Jindalee Stage B data (after Houminer *et al.*, 1996) containing E_s echo traces are also presented and clearly demonstrate the ease and effectiveness of the modelling and mapping techniques.

It should be noted that the effects of beam spreading, multiple hops, antenna patterns and absorption, on the BSI, have been ignored throughout this study.

10.2 The 4L ionospheric model

The 4L ionospheric model is simply an extension to the QPS ionospheric model in that it contains an extra layer, which represents the E_s. The 4L model consists of eight QP segments, one for each of the ionospheric layers and four QP segments representing the joining layers. The 4L model is an analytic ionospheric model, producing integrable equations when solving for the ray parameters. The joining segments which attach the E_s layer to the E layer as well as the equations for the E_s layer are given below.

The QP layers representing the E and E_s layers and the two joining segments, joining the E and E_s layers, are of the form:

$$N_E = a_E - b_E\left(1 - \frac{r_E}{r}\right)^2 \qquad \text{E layer}$$

$$N_{Uj} = a_{Uj} + b_{Uj}\left(1 - \frac{r_{Uj}}{r}\right)^2 \qquad \text{Upper joining segment}$$

$$N_{Lj} = a_{Lj} + b_{Lj}\left(1 - \frac{r_{Lj}}{r}\right)^2 \qquad \text{Lower joining segment}$$

$$N_{Es} = a_{Es} - b_{Es}\left(1 - \frac{r_{Es}}{r}\right)^2 \qquad E_s \text{ layer}$$

where

$$a = N_m$$

and

$$b = N_m \left(\frac{r_b}{y_m} \right)^2$$

where N_m is the maximum electron density occurring at the peak height r_m of the layer, r_b is the base height of the layer and y_m is the layer semi-thickness.

The electron concentration gradients in each of these segments are given by:

$$\frac{dN_E}{dr} = -\frac{2r_E b_E}{r^2}\left(1 - \frac{r_E}{r}\right)$$

$$\frac{dN_{Uj}}{dr} = \frac{2r_{Uj} b_{Uj}}{r^2}\left(1 - \frac{r_{Uj}}{r}\right)$$

$$\frac{dN_{Lj}}{dr} = \frac{2r_{Lj} b_{Lj}}{r^2}\left(1 - \frac{r_{Lj}}{r}\right)$$

$$\frac{dN_{Es}}{dr} = -\frac{2r_{Es} b_{Es}}{r^2}\left(1 - \frac{r_{Es}}{r}\right)$$

Assuming that the joining layer attaches smoothly to the other layers, the parameters of the joining layer can now be readily determined. The 4L model fits the joining layer, which is represented by two QP segments smoothly attached, between the E layer and E_s layer in the following manner.

Firstly, the joining segment is fitted to the E_s layer at a height, r_{jEs}, such that

$$r_{jEs} = r_{Es} + 0.9 y_m E_s.$$

The joining segment is given a minimum plasma frequency of:

$$a_j = 0.2 a_{Es}$$

where $a_j = a_{Uj} = a_{Lj}$.

The height of the minimum plasma frequency of the joining segment is given by:

$$r_j = \frac{r_{Es} b_{Es}\left(\frac{r_{Es}}{r_{jEs}} - 1\right)}{a_{Es} - a_j + b_{Es}\left(\frac{r_{Es}}{r_{jEs}} - 1\right)}$$

and

$$b_{Lj} = \frac{-r_{Es} b_{Es}\left(1 - \frac{r_{Es}}{r_{jEs}}\right)}{r_j\left(1 - \frac{r_j}{r_{jEs}}\right)}$$

where $r_j = r_{Uj} = r_{Lj}$.

The height at which the joining layer smoothly attaches to the E layer is given by:

$$r_c = \frac{r_E b_E\left(\frac{r_E}{r_j} - 1\right)}{a_E - a_j + b_E\left(\frac{r_E}{r_j} - 1\right)}$$

and

$$b_{Uj} = \frac{-r_E b_E\left(1 - \frac{r_E}{r_c}\right)}{r_j\left(1 - \frac{r_j}{r_c}\right)}.$$

Adding these equations to those of the QPS model produces the 4L model.

If $h_mE_s > h_mE$, a similar approach to that described above can be implemented or a four layer ionospheric model using the same criteria as the three layer, or QPS model can be used. Using the latter option there will be know valley regions between the layers, so if $h_mE < h_mE_s < h_mF_1$, and $f_oE_s > f_oF_1$ then the F_1 layer is ignored and the sporadic E layer fits directly to the F_2 layer. Both methods work well, producing realistic results.

10.3 The effects of E_s on synthesised BSI's using the 4L model

Figure 1 shows an example of a profile obtained using the 4L model with the following layer parameters.

f_oF_2 = 4.9 MHz	y_mF_2 = 125.0 km	h_mF_2 = 230.0 km
f_oF_1 = 4.0 MHz	y_mF_1 = 20.0 km	h_mF_1 = 142.6 km
f_oE = 3.12 MHz	y_mE = 11.0 km	h_mE = 102.0 km
f_oE_s = 4.0 MHz	y_mE_s = 1.0 km	h_mE_s = 90.0 km

A very narrow QP layer in Figure 1 represents the sporadic E. This sporadic E layer has a semi-thickness of 1 km, peak height of 90 km

and peak plasma frequency of 4 MHz. Figure 2 shows a blow up of the E_S layer. Clearly, a parabolic layer smoothly attached to the rest of the ionospheric profile represents the E_S. The valley region between the E and E_S consists of two smoothly attached QPS's.

Figure 3 represents a synthesised BSI determined using the QPS ionospheric model, which produced the profile in Figure 1, without the E_S layer.

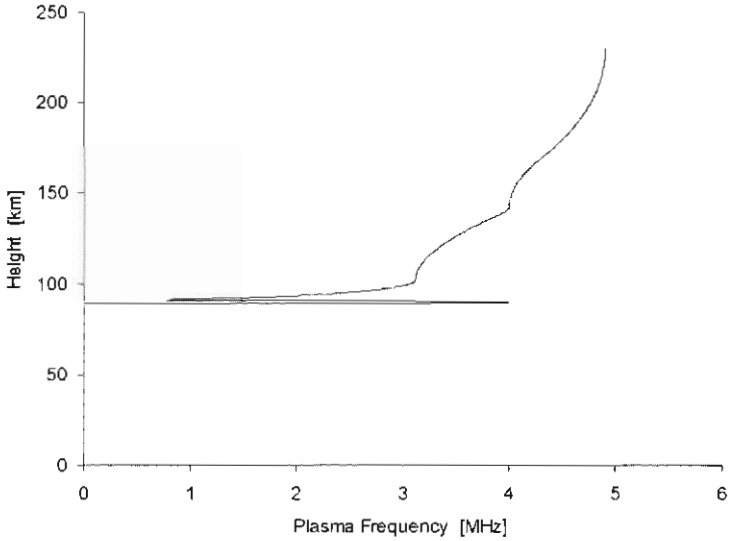


Figure 1. *An ionospheric profile containing a sporadic E layer.*

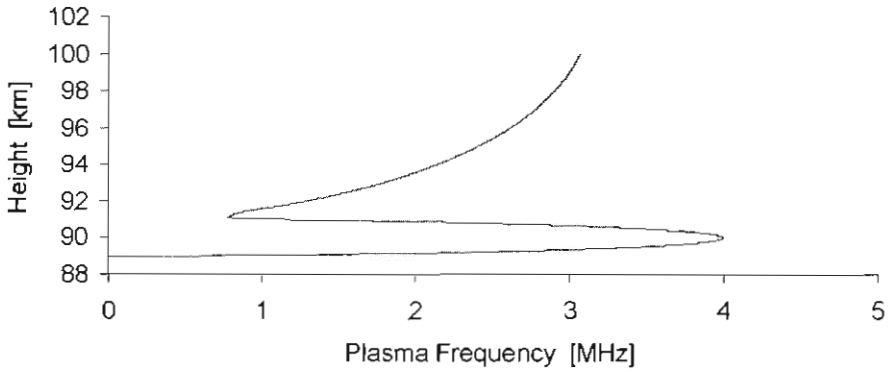


Figure 2. *A magnification of the sporadic E layer in Figure 1.*

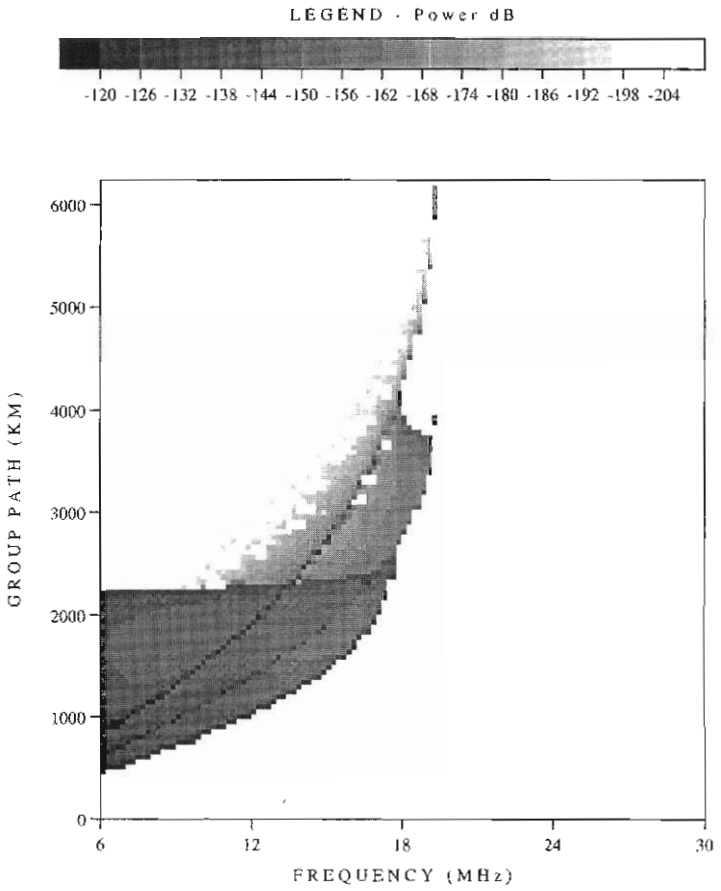


Figure 3. *Synthesised backscatter ionogram with no sporadic E.*

Figure 4 represents a synthesised BSI determined using the 4L ionospheric model, which produced the profile in Figure 1. Blanketing sporadic E was assumed in this case, i.e., there was no partial reflection by the E_s layer. Notice that in Figure 4 the F-layer trace contains ray paths reflected in the F layer where the plasma frequencies must have been greater than f_oE_s.

Figure 5 shows the ray take-off elevation as a function of both group path and frequency for the case of blanketing E_s. The elevation contours basically show where the sporadic E is relative to the transmitter location. For example, if there is a patch of E_s situated so only ray paths with elevation angles between 0 and 4 degrees intersect this cloud, then one would expect to obtain an E_s echo trace with the width of the lightest shaded contour in Figure 5. The heights, or altitudes, of the E_s clouds

occur in a relatively small range. If the approximate height of the E_s cloud is known, then using simple geometry the extent and distance from the transmitter/receiver location, of the E_s cloud can be mapped. If multiple traces of E_s appear on a BSI, then this approach can be used to map each cloud separately. The case of multiple echo traces on the BSI is not uncommon.

In this example a QP layer with a semi-thickness of 1 km is used to represent the E_s. The semi-thickness in these results has little affect on the E_s echo trace. Increasing the semi-thickness of the E_s to 10 km produces very little change in the shape of the E_s echo trace appearing on the BSI. The sporadic E behaves similarly to that of a spherical mirror. The only noticeable difference when increasing the semi-thickness of the E_s occurs

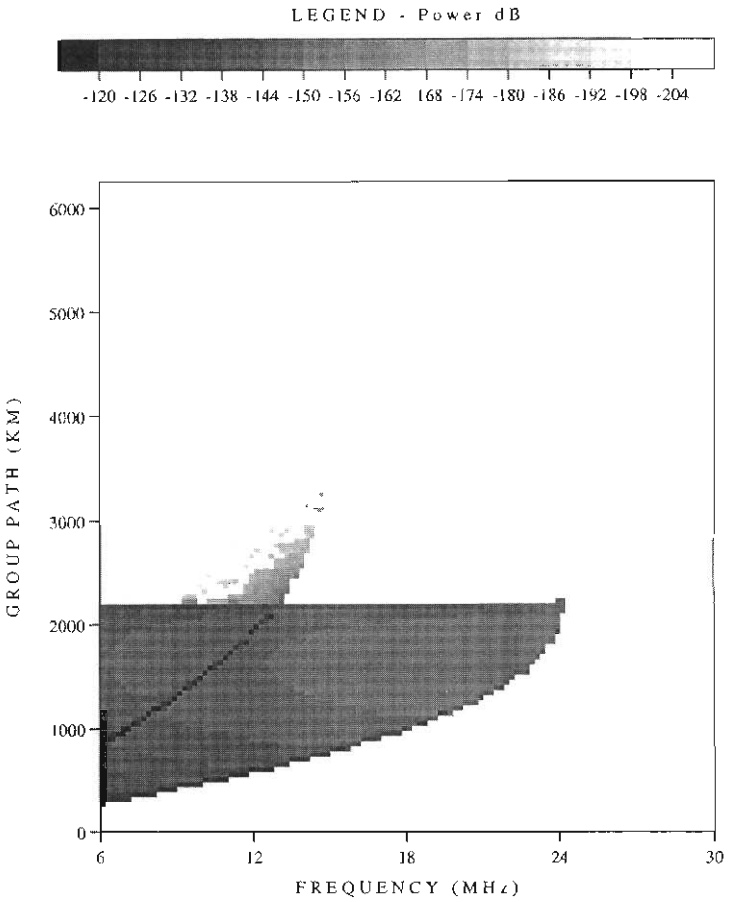


Figure 4. *Synthesised BSI with E_s, where f_oE_s = 4 MHz.*

near the peak operating frequency of the E_s echo trace where there is a slight upward bending in the E_s echo trace. This upward bending may be of some use in determining the actual thickness of the sporadic E cloud from real BSI's.

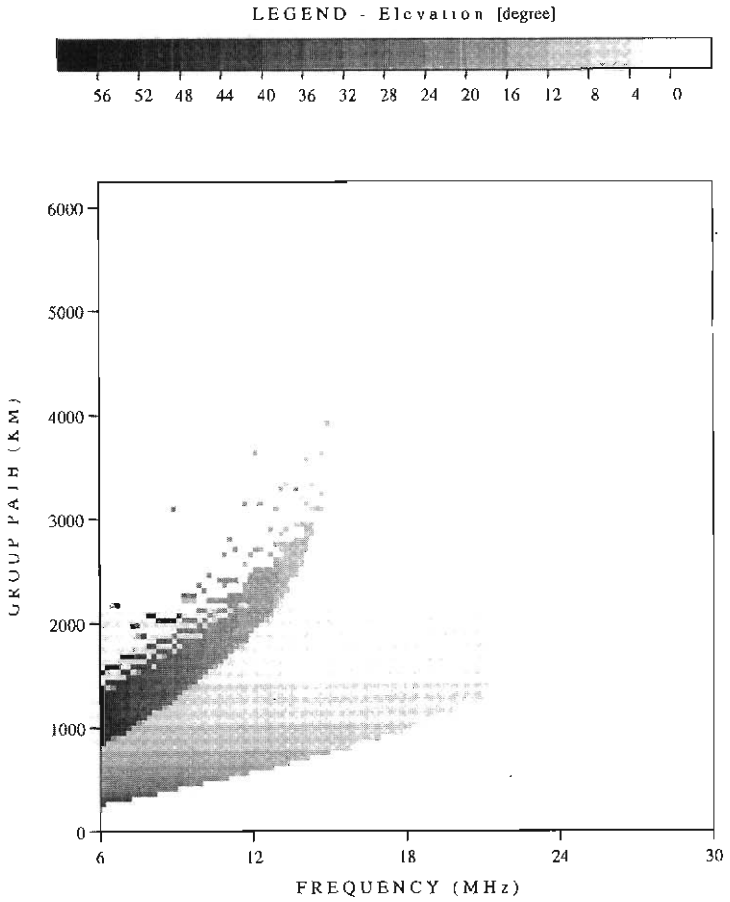


Figure 5. Ray take-off elevation as a function of both group path and frequency.

10.4 Modelling the E_s from the BSI

10.4.1 *Determining h_mE_s from a BSI when E_s is present over the entire coverage area*

Sporadic E occurs at altitudes between 90 and 130 km above the Earth's surface. Depending where you are positioned, this range can normally be reduced even further. It is more than likely that this parameter would be known to the frequency management system of over-the-horizon radar facilities and would not need to be determined. We have developed a

relatively simple method to calculate the h_mE_s for the special case of when the E_s is present over the entire coverage area. The minimum detectable elevation angle that the receiver can receive is also required in order to determine h_mE_s.

The h_mE_s can be determined from the BSI by assuming the E_s behaves in the manner of a spherical mirror. The synthesised BSI in Figure 4 is used as an example to illustrate the procedure to determine h_mE_s.

Firstly, the maximum group path from the leading edge of the BSI is determined. In this case the maximum group path is approximately 2250 km.

Secondly, assuming the E_s behaves as a good mirror model, then the following equation, which determines the group path, for the ray from the transmitter to the base of the E_s layer, can be rearranged as follows to produce a simple equation to calculate h_mE_s.

$$P' = 2\left(\sqrt{\left(r_b^2 - r_o^2 \cos^2 \beta_o\right)} - r_o \sin\beta_o\right) \qquad (1)$$

where β_o is the smallest angle of elevation that the receiver can detect from the E_s layer,

r_o represents the radius of the Earth, and

r_b represents the radius from the centre of the Earth to the E_s cloud, $h_mE_s + r_o$.

Then Equation (1) can be rewritten as:

$$r_b = \sqrt{\left(\frac{P'}{2} + r_o \sin\beta_o\right)^2 + r_o^2 \cos^2 \beta_o} \quad .$$

Depending on the antennas and their location, in practice, $\beta_o = 2°–3°$, at best. However, the synthesised BSI in Figure 4 displays ray paths with elevation angle beginning at $\beta_o = 0.1°$. From the BSI in Figure 4, $P' \approx 2150$ km.

Then $r_b = h_mE_s + r_o \Rightarrow h_mE_s \approx 92$ km which is close to the correct value of 90 km for the peak height of the E_s layer used in producing the synthesised BSI in Figure 4. Using a more accurate estimate of P' would produce a more accurate value of h_mE_s. In the following sections, equations for f_oE_s, the length of the E_s cloud and the distance the E_s cloud is away from the transmitter/receiver location, are presented and show that they are in terms of r_b and not h_mE_s. This means that a high accuracy in the value of h_mE_s is not crucial for the modelling and mapping of E_s and depending on the desired accuracy in the modelling and mapping required an educated guess of h_mE_s is in practice normally sufficient.

10.4.2 *Determining f_oE_s from a BSI*

The peak plasma frequency should be determined for each of the E_s clouds. The equation used to determine the peak plasma frequency of an E_s cloud, f_oE_s, is derived as follows:

At reflection, we know $\mu^2r^2 - r^2_o\cos^2\beta_o = 0$

where μ represents the refractive index,
r represents the height of the ray in this case $r = h_mE_s + r_o$ and
r_o represents the radius of the Earth ($r_o = 6370$ km).

In the absence of an imposed magnetic field and collisions, the refractive index is given by

$$\mu^2 = 1 - \frac{f_N^2}{f^2}$$

where f_N represents the plasma frequency (at height $h_mE_s = f_oE_s$) and
f represents the operating frequency or wave frequency.

Then

$$f_N = \sqrt{f^2\left(1 - \frac{r_o^2\cos^2\beta_o}{r^2}\right)} \qquad (2)$$

and from the BSI the highest wave frequency, f, reflected from the E_s can easily be measured. The parameter β_o is the lowest angle of elevation for a ray that intersects the E_s cloud. By substituting these values into Equation (2), f_N can be determined.

As an example, the f_oE_s for the BSI in Figure 4 will be determined.

Firstly, from the BSI, $f = 24$ MHz and $r = 90 + r_o$.

The lowest detectable elevation angle in this case is 0.1 degree.

Thus, $f_N = f_oE_s = 4$ MHz which is the correct value.

10.5 Mapping E_s

In order to map E_s from a BSI one must determine the distance the transmitter/receiver location is away from the E_s cloud. The extent of the E_s cloud in the transmitter/receiver also needs to be determined from the BSI. Sometimes more than one E_s echo trace is present on the BSI where each echo corresponds to a cloud of E_s. It is highly likely that the f_oE_s differ from one cloud to the next. The following procedure should be performed on each cloud separately.

Firstly, calculate the maximum and minimum elevation angles that are reflected from the E_s clouds. This is achieved by reading the maximum and minimum group path of the E_s echo trace and rearranging Equation (1) for a spherical mirror model, as follows:

$$\beta_{min} = \sin^{-1}\left(\frac{r_b^2 - \frac{P'^2_{max}}{4} - r_o^2}{P'_{max} r_o}\right)$$

and

$$\beta_{max} = \sin^{-1}\left(\frac{r_b^2 - \frac{P'^2_{max}}{4} - r_o^2}{P'_{min} r_o}\right)$$

where β_{max} is the maximum elevation angle of ray paths which transverse the E_s cloud,

β_{min} is the minimum elevation angle of ray paths which transverse the E_s cloud,

P'_{max} is the maximum group path of the E_s echo read from the BSI,

P'_{min} is the minimum group path of the E_s echo read from the BSI,

r_o represents the radius of the Earth, and

r_b represents the radius from the centre of the Earth to the E_s cloud, $h_mE_s + r_o$.

The extent or length, L, of the E_s cloud in the transmitter/receiver direction is given by this simple equation:

$$L = r_b(\psi 1 - \psi 2)$$

where

$$\Psi 1 = \frac{\pi}{2} - \alpha_{max} - \beta_{min}$$

$$\Psi 2 = \frac{\pi}{2} - \alpha_{min} - \beta_{max}$$

and

$$\alpha_{max} = \sin^{-1}\left(\frac{r_o}{r_b}\sin\left(\frac{\pi}{2} + \beta_{min}\right)\right)$$

$$\alpha_{min} = \sin^{-1}\left(\frac{r_o}{r_b}\sin\left(\frac{\pi}{2} + \beta_{max}\right)\right)$$

where α is the angle of incidence at the base of the ionosphere (E_s patch). The parameter ψ is the angle between the transmitter and the E_s cloud relative to the centre of the Earth, and $\psi 1$ and $\psi 2$ are the angles of the ray paths with the lowest and highest elevation angles, which intersect the E_s cloud.

The equation below determines the ground distance, D, in kilometers, from the transmitter/receiver location to the position on the ground

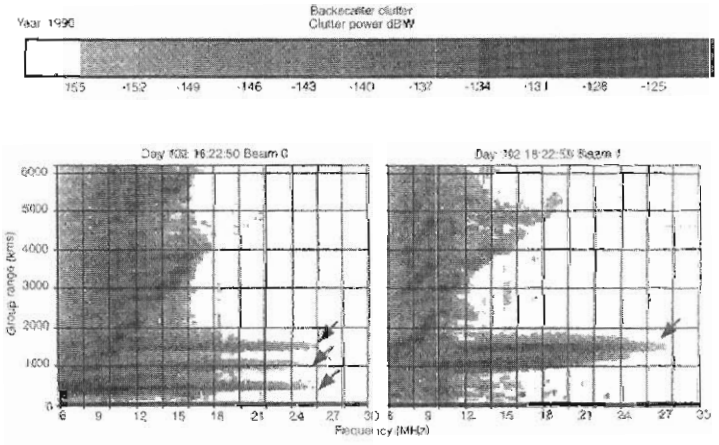
directly below the edge of the E_s cloud nearest the transmitter/receiver location:

$$D = r_o\psi 2 = r_o\left(\cos^{-1}\left(\frac{r_o}{r_b}\cos\beta_{\max}\right) - \beta_{\max}\right)$$

The actual location of the E_s cloud in terms of geographic or geomagnetic coordinates can now be readily determined.

10.6 Testing the technique on real BSI's

Two examples are presented here which clearly demonstrate the ease and effectiveness of the modelling and mapping techniques developed in this study. Figures 6 and 7 show the two Backscatter ionograms from the Jindalee Stage B data (after Houminer *et al.*, 1996) which we will attempt to model and map. They are both nighttime cases and clearly show a weak background ionosphere.



Figures 6 and 7. *Examples of Jindalee backscatter ionograms with pronounced sporadic E echoes (arrows), (after Houminer et al., 1996). Times in UT.*

10.6.1 Example 1 modelling and mapping the BSI in Figure 6

The BSI in Figure 6 shows three distinct E_s echo traces at group paths in the range 400–1700 km, indicating three distinct E_s clouds were present at this time. The three traces are quite narrow indicating that the patches of E_s were relatively small. The three traces extend out to nearly 27 MHz indicating that they are quite dense.

In order to model the E_s, one must firstly, measure the maximum operating frequency, and the maximum and minimum group paths of each of the E_s echo traces on the BSI in Figure 6. The results in Table 1 display these measured values. The maximum operating frequency is used to calculate the maximum plasma frequency of each of the three E_s clouds. The maximum and minimum group path measurements are used to calculate the extent of each of the E_s clouds. Let E_s1 represent the E_s cloud closest to the receiver location, represented by the E_s echo trace with the smallest group paths. Let E_s2 represent the cloud second closest to the receiver and let E_s3 represent the E_s cloud furthest from the receiver location.

Table 1.

E_s Cloud	Max. Frequency (MHz)	Min. Group Path (km)	Max. Group Path (km)
E_s1	25.7	350	600
E_s2	26	900	1150
E_s3	26.5	1450	1700

It is assumed in this case that h_mE_s for each of the patches is 90 km, again noting that E_s generally occurs at altitudes between 90 and 130 km.

Secondly, the maximum and minimum elevation angles that are reflected from the E_s clouds and the extent or length, L, of the E_s cloud in the direction of the BSI as well as the ground distance, D, in kilometers, from the transmitter/receiver location to the E_s cloud are calculated using the equations described earlier, and are presented in Table 2.

The actual location of the E_s cloud in terms of geographic or geomagnetic coordinates can be readily determined.

Table 2.

E_s Cloud 127	Min. β_0 (degrees)	Max. β_0 (degrees)	Length of Cloud L (km)	Distance D (km)
E_s1	16.2	30.2	137	149
E_s2	6.5	9.5	128	438
E_s3	2.3	3.9	127	715

Thirdly, the peak plasma frequency for each of the E_s clouds was calculated using the equations derived earlier, along with the maximum operating frequencies of the ray paths which were reflected from each of

the E_s echo traces. The peak plasma frequencies for each of the E_s clouds are shown in Table 3.

Table 3.

E_s cloud	Max. Frequency (MHz)	F_0E_s (MHz)
E_s1	25.7	8.25
E_s2	26	5.2
E_s3	26.5	4.53

Then placing these parameters into the new 4L analytic ionospheric ray tracing model, the synthesised BSI in Figure 8 was produced. The E_s echo traces, in this BSI, match well to the actual E_s echo traces in the BSI shown in Figure 6.

There is very little curvature in the E_s echo traces in Figures 6 and 7 indicating very thin E_s clouds. Thus, a semi-thickness of 1 km was used in synthesising the E_s clouds.

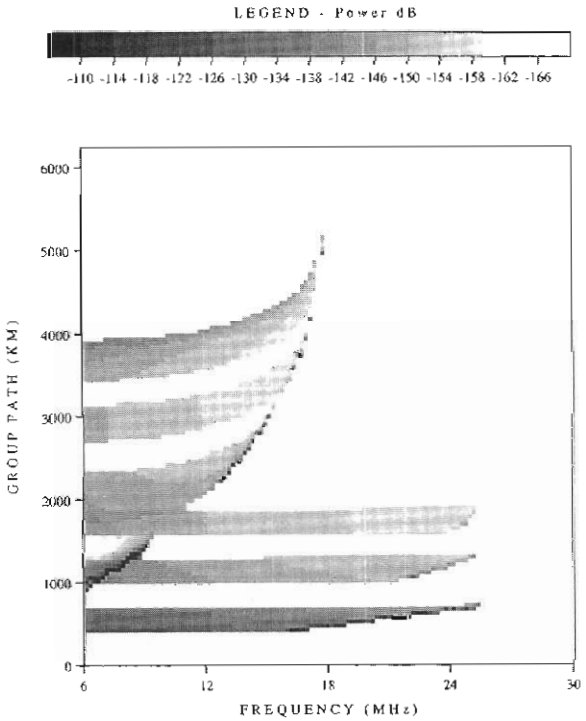


Figure 8. *Synthesised BSI with multiple clouds, or patches of E_s.*

Figure 8 shows our first attempt at modelling the traces of the BSI in Figure 6, where we have assumed blanketing patches of sporadic E.

Figure 9 shows the synthesised BSI when assuming only partial reflection from the E_s clouds. This synthesised BSI matches very well to that shown in Figure 6. In this case it is assumed that all ray paths are only partially reflected from the E_s clouds. The method used is designed so that there is about a 20 dB transmission loss for ray paths reflected from $(h_mE_s - 0.2)$ to h_mE_s and a 7 dB transmission loss for ray paths reflected from the base of the E_s clouds to $(h_mE_s - 0.2)$. An antenna gain of 10 dB was also included in the synthesised BSI shown in Figure 9.

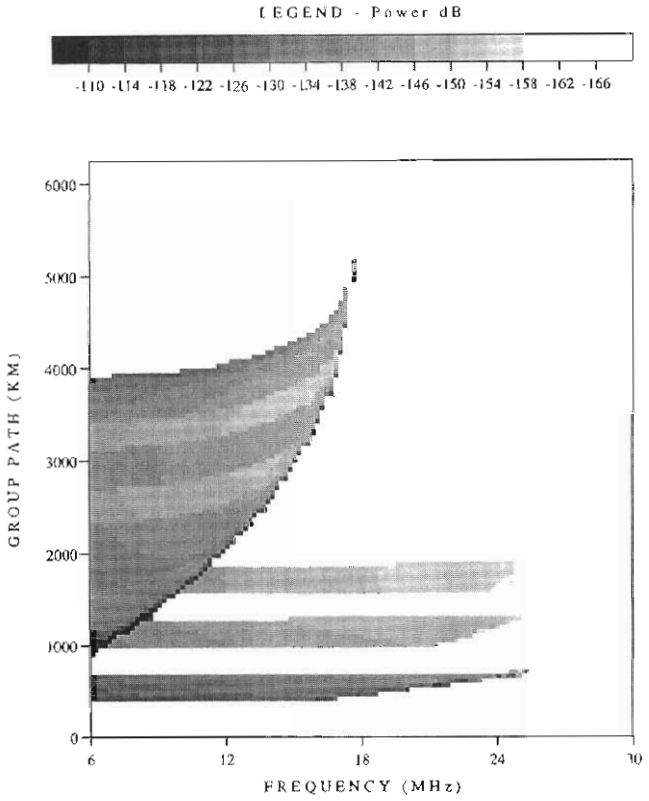


Figure 9. *Synthesised BSI having multiple traces of E_s with partial reflection from the E_s clouds.*

10.6.2 Example 2 modelling and mapping the BSI in Figure 7

The BSI in Figure 7 has one well-defined E_s echo trace in the 900–1900 km group range. This means that only one large E_s cloud is present. The trace appears to have a stronger region of backscatter clutter at a group range of 1600 km. This stronger region is due to a dense patch

of ionisation within the E_s cloud. This can easily be modelled using the technique used previously in modelling the BSI in Figure 6. However, in this section the E_s echo trace is modelled as a single spherically stratified E_s cloud.

The maximum operating frequency of the ray paths reflected from the E_s cloud is 27.5 MHz. The maximum group range is approximately 1700 km. Since we do not have the received angles-of-arrival of these ray paths, let us assume that the lowest detectable elevation of the antenna

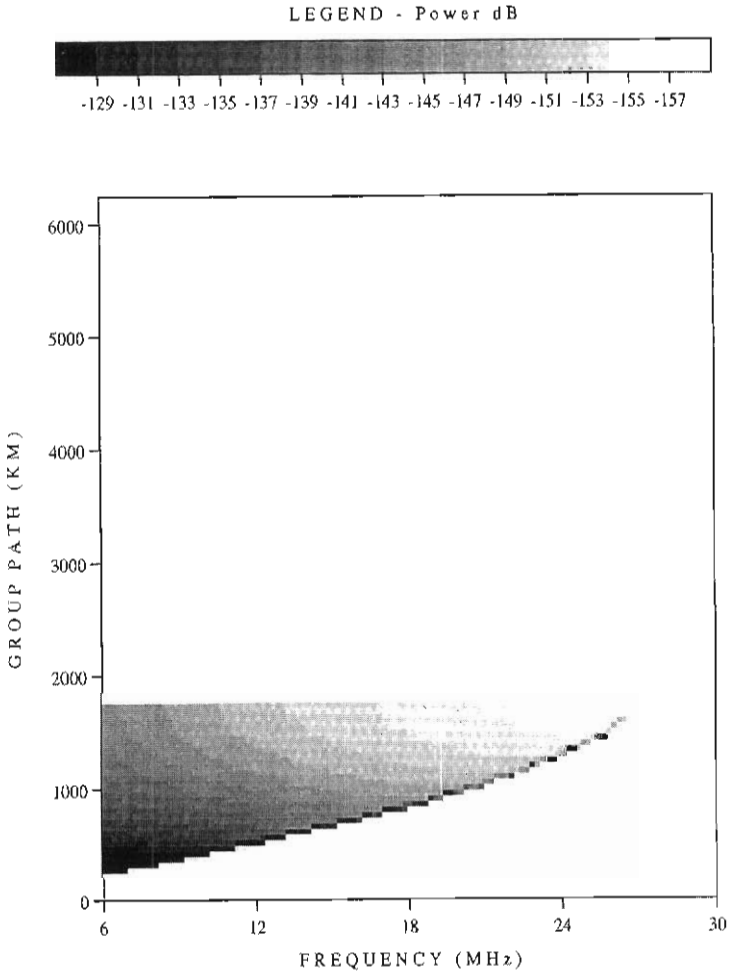


Figure 10. *Synthesised BSI with blanketing E_s.*

array is 3°. Then h_mE_s and f_oE_s are determined using Equations (1) and (2) and found to be 90 km and 5 MHz respectively.

Then placing these parameters into the new 4L analytic ray tracing model, the synthesised BSI shown in Figure 10 resulted. The shape of the E_s echo trace in this synthesised BSI matches well with the actual E_s echo trace in the BSI shown in Figure 7.

Applying the same conditions of partial reflection, to those that were used in modelling Figure 6, the synthesised BSI in Figure 11 was obtained, which compares reasonably well to the real BSI in Figure 7. Knowing the actual antenna gains would produce even better results.

Not only is it now possible to model the E_s, it is also possible to estimate the size and distance from the transmitter/receiver location of the E_s clouds.

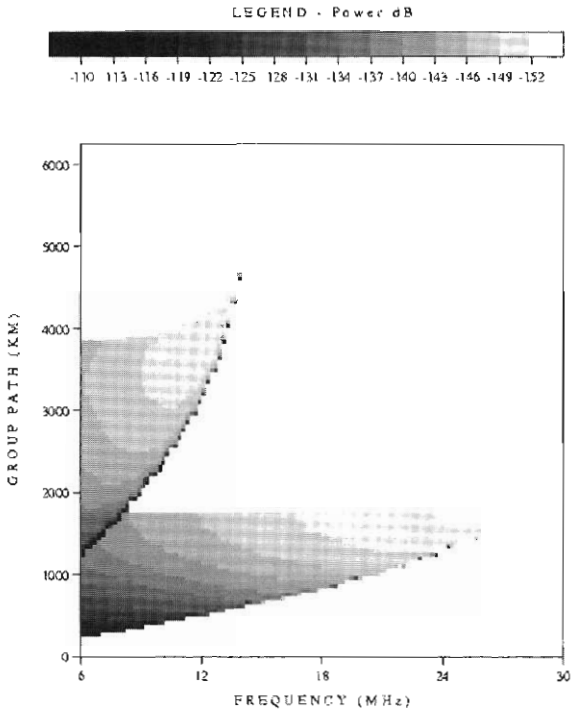


Figure 11. *Synthesised BSI with partial reflection from the E_s.*

10.7 Summary

The new 4L ionospheric model which is simply an extension of the QPS model, in that, another spherically stratified layer is added at the base of the ionosphere to represent Es, produces realistic results.

The equations, determining the distance of the E_s cloud from the transmitter/receiver location as well as the peak plasma frequency and the extent of the E_s cloud, produce encouraging results. The E_s echo traces on the BSI's compared reasonably well with the synthesised BSI's, where the synthesised BSI's were determined using the 4L ionospheric model.

The techniques of both mapping and modelling E_s, developed here, are relatively simple, require few input parameters and may be useful in the frequency management system of over-the-horizon radar facilities.

References

Dyson, P.L. and Bennett, J.A. (1988). A model of the electron concentration in the ionosphere and its application to oblique propagation studies. *Journal of Atmospheric and Terrestrial Physics,* 50: 251–262.

Houminer, Z., Russell, C.J., Dyson, P.L. and Bennett, J.A. (1996). Study of sporadic-E clouds by backscatter ionograms. *Annales Geophysicae,* 14: 1060–1065.

Norman, R.J., Dyson, P.L. and Bennett, J.A. (1998). Modeling and mapping sporadic E using backscatter ionograms. *Report to the Technical Steering Group, Jindalee Project.* Telstra Corporation. p. 21.

11. IMAGING RIOMETER OBSERVATIONS OF ABSORPTION PATCHES ASSOCIATED WITH MAGNETIC IMPULSE EVENTS

M.B. Terkildsen[1,2], B.J. Fraser[1], F.W. Menk[1] and R.J. Morris[2]

[1] Department of Physics
CRC for Satellite Systems
University of Newcastle
Newcastle NSW 2308 Australia

[2] Atmospheric and Space Physics
Australian Antarctic Division
Kingston Tasmania 7050 Australia
(email: michael.terkildsen@aad.gov.au)

Abstract

Current research utilises the recently installed Southern Hemisphere Imaging Riometer (SHIRE) located at Australia's Davis Antarctic station to study riometer signatures associated with cusp-latitude magnetic impulse events (MIE's) detected in co-located fluxgate magnetometer data. Such observations enable further refinement of the characterisation of MIE's by observing the spatial scale, morphology and dynamics of associated patches of locally enhanced Cosmic Noise Absorption (CNA). The multiple narrow beams which comprise the field-of-view of an imaging riometer allow us to obtain two-dimensional images of the localised absorption patches as they evolve over time, facilitating study of the structure and motion of MIE-associated absorption events. Of those dayside MIE's with a detectable riometer signature, two distinct morphologies have emerged. The first of these has many characteristics of the morning sector absorption spikes reported by Stauning and Rosenberg (1996), yet with a previously unreported spatial and temporal bipolarity in the riometer signature. The second and more common signature observed in association with magnetic impulses is a widespread increase in absorption of spatial scale larger than the 200 x 200 km^2 field-of-view of the imaging riometer, but with generally smaller amplitude and consistent antisunward and magnetic poleward convection. In this paper we present two case studies illustrating both types of events.

11.1 Introduction

Magnetic Impulse Events (MIE's) observed in high-latitude magnetograms are perturbations in the geomagnetic field which manifest themselves as a half to a full cycle of a 300–800 s period

pulsation. Such events have previously been explained in terms of the ionospheric response to transient magnetic reconnection at the magnetopause, commonly termed Flux Transfer Events (FTE's) (Russell and Elphic, 1979; Glassmeier, 1984; Lanzerotti *et al.*, 1986). However a number of other magnetospheric phenomena have also been suggested as mechanisms likely to invoke a comparable ionospheric response. Such mechanisms include impulsive penetration of plasma into the magnetosphere (Heikkila *et al.*, 1989); abrupt changes in the solar wind dynamic pressure (Sibeck, 1989, 1990); and a Kelvin–Helmholtz instability acting at the low-latitude boundary layers (McHenry *et al.*, 1990). The investigation of these processes is a key element of high-latitude studies, as they provide insight into the mechanisms by which solar wind energy is coupled into the Earth's magnetosphere and subsequently transferred to ionospheric altitudes.

It is generally accepted that communication to the ionosphere of a magnetospheric disturbance such as a FTE, is through intermediate field aligned currents which close in the ionospheric E-region (Southwood, 1987; McHenry and Clauer, 1987). When reconnection commences at the magnetopause, solar wind plasma is linked to the ionosphere via convecting flux tubes which carry these field aligned currents. The exact nature of the field aligned current systems, and their closure currents in the ionosphere, is not well understood. However, the characteristic magnetic perturbations generated in response to the Hall component of the closure current have been extensively modelled, most recently by Chaston *et al.* (1993), Zhu *et al.* (1997), and Zhu *et al.* (1999).

Further studies have shown that in addition to a characteristic magnetic perturbation, there is commonly a concurrent signature in co-located riometers. A study by Korotova *et al.* (1997) using magnetometer and riometer data from the South Pole (-74.2° MLAT), suggests up to 80% of those impulsive magnetic events which are detected have an associated riometer signature. Two possible mechanisms may explain the enhanced radiowave absorption observed in association with magnetic impulses. The first is thought to be the most commonly observed type of absorption event and is due to enhanced ionisation resulting from the precipitation of energetic electrons. Such energetic electrons (with typical energies of 30–300 keV) originate from populations of quasi-trapped substorm produced electrons drifting around to the dayside. Their precipitation results in the enhancement of D-region (60–90 km) electron densities that subsequently enhance the level of radiowave absorption in the ionosphere. Alternatively, the large horizontal electric fields associated with currents in the ionosphere can raise the electron temperature of the intervening plasma sufficiently through non-deviative collisional heating. Hence, significant enhancements of E-region electron collision frequency may occur (typically at ~ 120 km altitude). Such an enhancement of electron–neutral collision frequency in a region of sufficiently high localised electron density (such as the daytime profile) can result in a cosmic noise absorption (CNA) enhancement detectable by a riometer (Stauning, 1984; Stauning *et al.*, 1985).

The Korotova *et al.* (1997) study, however, was made using a broadbeam riometer which integrates the received signals from a wide field-of-view and thus provides no information on the spatial scale or motion of the absorption event. In addition, a broadbeam riometer cannot distinguish spatial variations in cosmic noise absorption levels from temporal variations. On the other hand, observations with an imaging riometer, with its multiple narrow beams and superior spatial resolution, allows us to produce two dimensional 'images' of the patches of enhanced CNA as they evolve over time. Thus observations can be made of the spatial scale, morphology and dynamics of MIE–associated absorption patches. This enables further refinement of the characterisation of magnetic impulse events by assisting in distinguishing events with comparable magnetic signatures such as FTE's, precipitation spike events (Stauning and Rosenberg, 1996), and Travelling Convection Vortices (TCV's) (Glassmeier, 1992).

The motion of the ionospheric signatures of these events is particularly characteristic. TCV's propagate tailward typically with speeds of 2–10 km s^{-1}, with occurrence peaking at about 0900 and 1400 MLT. There is very little poleward motion of these signatures (Glassmeier, 1992). Alternatively, FTE's, or the ionospheric signatures of reconnection as modelled by Chaston *et al.* (1993), show a significant poleward component of motion of the footprint of the reconnected flux tube and an east–west component of motion that is dependent on IMF B_y. Typical velocities through the ionosphere are 0.53 km s^{-1} with occurrence peaking within a few hours about local magnetic noon.

The imaging riometer used to study the CNA signatures of cusp latitude impulsive magnetic events was the recently installed Southern Hemisphere Imaging Riometer (SHIRE) located at Australia's Davis Antarctic station. Events were initially identified using co-located three component fluxgate magnetometer data.

11.2 The southern hemisphere imaging riometer experiment

Over the 1996–1997 Austral summer, installation of the SHIRE imaging riometer at Davis, Antarctica (-74.5°S magnetic latitude) was completed. The project is a collaboration between the Space Physics Group at the University of Newcastle, the Institute for Physical Science and Technology (IPST) at the University of Maryland, and the Atmospheric and Space Physics group at the Australian Antarctic Division.

The SHIRE imaging riometer (see Figure 1), one of the Imaging Riometer for Ionospheric Studies (IRIS) systems (Detrick and Rosenberg, 1990), consists of a phased array of circularly polarised antennas in a square 8 × 8 element configuration. The array has its major axes aligned with the compass magnetic NS and EW directions with half-wavelength spacing between individual antenna elements. Signals received by each of the 64 crossed dipole ('turnstile') antennas are fed into an eight port Butler matrix which is used as the beam forming device (Butler and

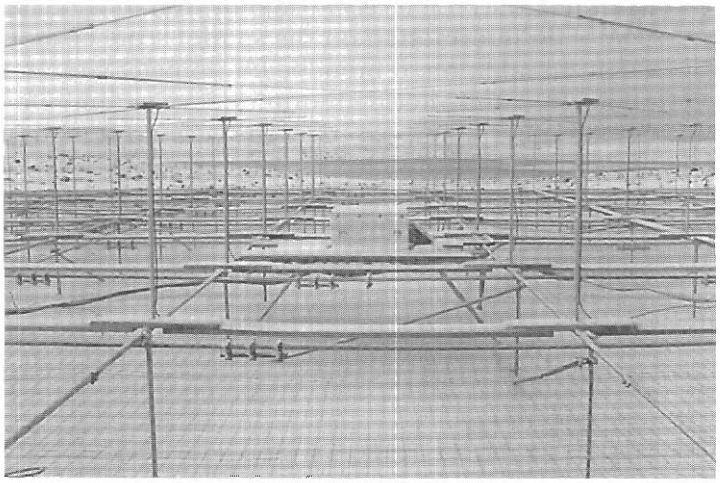


Figure 1. *Antenna array of the Southern Hemisphere Imaging Riometer Experiment (SHIRE), located at Davis, Antarctica (-74.5° MLAT).*

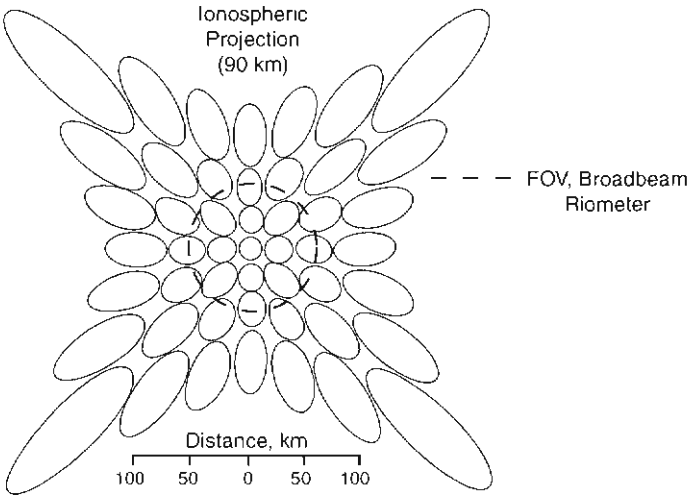


Figure 2. *Projection of the -3 dB contours of the 49 IRIS beams onto the ionosphere at 90 km altitude. The dashed circle in the centre represents the nominal (-3 dB) viewing area of a conventional broad-beam antenna (from Detrick and Rosenberg, 1990).*

Lowe, 1961). The standard Butler matrix has been modified to produce a zenith pointing beam resulting in the sacrifice of one of the output ports of the matrix. The result of such phasing is the formation of 49 conical beams each with a -3 dB beamwidth of approximately 13° and oriented in a unique direction forming a network of partially overlapping beams. Figure 2 shows the horizontal planar cross-section at ionospheric altitudes (assumed to be 90 km) of the -3 dB power level for all beams of the IRIS system. This is then the effective field-of-view of the instrument at 90 km altitude, which is the assumed average altitude for ionospheric absorption of cosmic radio waves (Detrick and Rosenberg, 1990).

SHIRE operates at 38.2 MHz with a 250 KHz bandwidth, and logs the 49 channels continuously at 1 Hz by sweeping sequentially across each row of detectors. The data logged from each beam is normalised to remove the diurnal variation in received cosmic noise power, and can be viewed as individual time series plots in a similar manner to the output of a broadbeam riometer, or as a time series of two dimensional 'images' of patches of enhanced CNA. The multiple narrow beam configuration of SHIRE also results in a wider area coverage (~ 200 x 200 km^2 at 90 km altitude), at significantly enhanced temporal and spatial resolution (1 s and up to 20 km, respectively).

SHIRE is situated at a magnetic latitude of -74.5° and passes through local magnetic noon at approximately 0950 UT. Such a location generally positions the riometer just equatorward of the footprint of the cusp, near the last closed magnetic field line. At this latitude, magnetic field lines regularly map to the outermost regions of the magnetosphere, such as the dayside magnetopause where the dynamic interactions of the solar wind and the Earth's magnetosphere take place.

11.3 Case studies

The selection criteria used for the detection of MIE's in fluxgate magnetometer data was based on that of Lanzerotti *et al.* (1991). The identification routine stipulate the magnetic deflection satisfy a minimum amplitude criteria (taken as $\Delta Z > 40$ nT, and ΔX, $\Delta Y > 20$ nT), and that it is of a temporal duration of 300–900 s. It also requires the deflection to be distinguishable from noisy magnetic records by limiting the number of extrema in a temporal window surrounding the event, and requiring ΔZ to be above background fluctuations. Although a necessary selection criteria, the requirement of an event free from magnetic noise is not ideal as it effectively removes events occurring during magnetically active times when Davis moves toward open field lines, and to a lesser extent events around magnetic noon where average dayside magnetic activity peaks.

In a preliminary investigation, a number of impulsive events were identified in 1997 fluxgate magnetometer data recorded at Davis, and their associated riometer signatures were studied. Among those events with a detectable riometer signature, some distinct morphologies have been observed, namely short duration localised absorption events

detected in only a small number of beams of the array and often showing a bipolar structure, and secondly (and more commonly) a widespread increase in CNA observed in all beams of the array indicating an event of spatial scale larger than the 200 × 200 km^2 field-of-view of the instrument.

We now examine two representative examples satisfying these exclusive criteria.

11.3.1 Event of 0730 UT, 18 September 1997

Figure 3 shows an event of the first type – an MIE with an associated riometer signature consisting of a short duration spike. A distinct event in the magnetogram, this magnetic impulse occurred in the local morning at Davis (~ 0930 MLT) during a period of moderate magnetic activity ($K_p = 3^+$). The figure shows the relative timing and magnitude of the fluxgate magnetometer signature (bottom three traces) and the riometer signature observed in the (1, 0) beam of the SHIRE array (top trace). Taking the time of event occurrence as the point at which the signal reaches 90% of it's peak value, the magnetic signature and the riometer signature – a bipolar modulation of local radiowave absorption levels of considerably shorter period – can be seen to occur simultaneously when the relative fields-of-view of the instruments are taken into consideration.

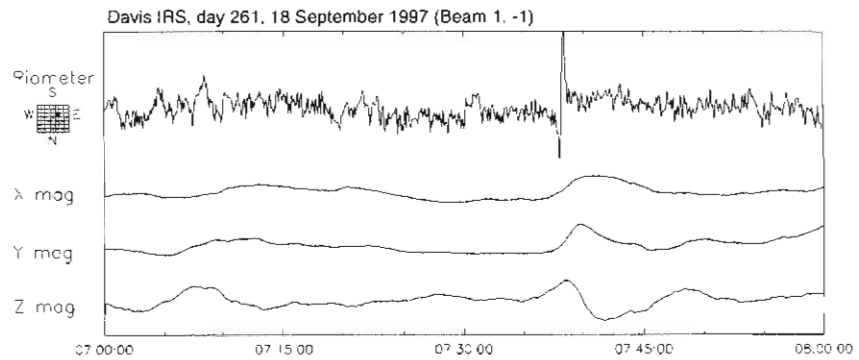


Figure 3. *Time series of received signal intensity in one beam of the SHIRE array, and the associated magnetic signature as measured by a co-located three-component fluxgate magnetometer on 18 September 1997. A bipolar spike in absorption can be seen associated with the longer period magnetic impulse at 0738 UT. A positive deflection in the riometer trace indicates an increase in radiowave absorption level.*

To observe the spatial and temporal structure of this absorption event, all 49 beams of the array were combined using an interpolation routine to produce a time series of absorption 'images' (see Figure 4). Each square in this figure represents a 200 x 200 km^2 area of the ionosphere with the time series progressing left to right, top to bottom. Absorption intensities are color coded according to the scale on the bottom of the figure. The absorption scale is deliberately centred around an average taken to be the background level of absorption and set to zero so that both positive and negative deflections in absorption levels can be observed.

The convection or phase velocity of the event can be estimated through a cross-correlation technique along specific rows of the array using the data shown in Figure 4. This technique assumes the event has the form of a steadily propagating front resulting in constant time delays between detection in successive beams along each row of the array. After converting to a geomagnetic coordinate system, this yields a westward (i.e. antisunward) phase velocity of ~ 600–700 m s^{-1}.

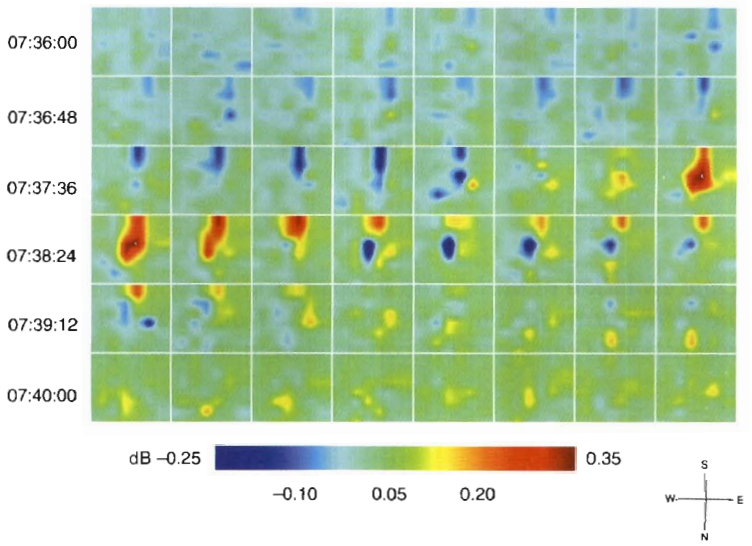


Figure 4. *Time series of absorption images (L to R, T to B) for the interval 0736–0740 UT on 18 September 1997 showing the progression of a MIE associated absorption event. Each image represents an ~ 200 x 200 km^2 view of the ionosphere, with an ~ 6 s time difference between adjacent panels. The orientation of the images is indicated at lower right. The absorption scale is centred on the average taken as zero so that both positive and negative deflections in absorption intensity may be observed.*

Figure 5 shows the evolution of absorption intensities over time along one column and one row of the SHIRE array. Again the absorption scale is centred around an average taken as zero, and a clear westward progression of the absorption feature may be observed. Plotting data from other rows in the same manner reveals a substantial (~ 200 m s^{-1}) southward (i.e. poleward) component of motion of the main peak in intensity, in addition to the predominant westward progression.

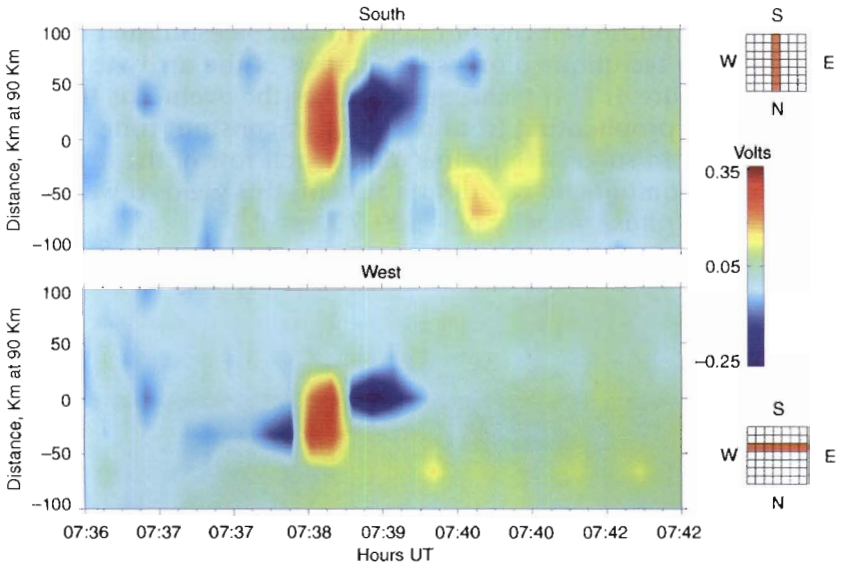


Figure 5. *Time evolution of absorption intensities along one row and one column of the SHIRE array for the event at 0738 UT on 18 September 1997. A westward progression of the signature can be seen in the lower panel.*

Figure 6 shows the time series absorption data from the full SHIRE array with the frames oriented on the page in the same order that the beams project onto the ionosphere as viewed from below. A positive deflection in this figure indicates an increase in the level of radiowave absorption (in dB). At any one instant, the absorption spike can be seen to be restricted to a small number of beams. The size of the event is no more than 70 (40) km in N–S (E–W) extent, and 50–150 s in total duration.

IMF and solar wind conditions (as measured by the WIND satellite located 88 R_E upstream) during and immediately prior to the impulse event were relatively stable. IMF B_x, B_y, and B_z were all strongly negative during the event (-4 nT, -10 nT, and -7 nT respectively), circumstances which favour the occurrence of FTE's.

In summary for this event the features of note are:

(i) A spatially and temporally restricted absorption event, no more than 70 (40) km in N–S (E–W) extent, and 50–150 s in total duration.

(ii) Bipolar structure in absorption signature.

(iii) Convection of the event at ~ 600–700 m s^{-1} in a predominantly westward (antisunward) direction, and ~ 200 m s^{-1} poleward.

(iv) Event properties consistent with those of an FTE.

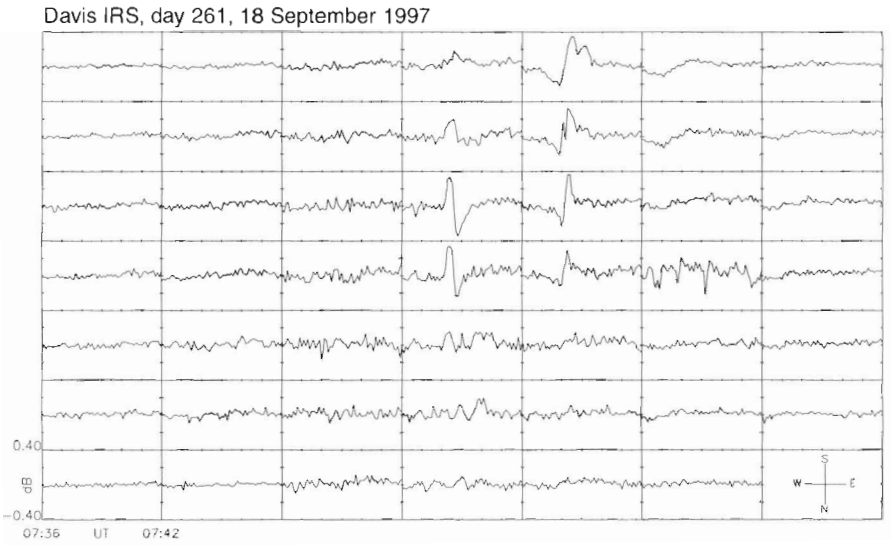


Figure 6. *Absorption intensities (in dB) in all 49 beams of the SHIRE array during the interval 0736–0742 UT on 18 September 1997. A positive deflection indicates an increase in radiowave absorption levels in this figure. The frames are arranged in the order that the beams project onto the ionosphere as viewed from below. Magnetic south is to the top of the page.*

11.3.2 Event of 0820 UT, 22 August 1997

The second event presented here (see Figure 7) also occurred pre-noon (~ 1020 MLT), and was observed by the Davis magnetometer as a large impulsive bipolar deflection in the Z component of the magnetic field. Such a magnetic perturbation is indicative of dual Hall current loops convecting overhead in the ionosphere, consistent with interpretation of the event in terms of a TCV (Friis-Christensen *et al.*, 1988). However, it must be noted there is no unambiguous relation between the magnetic perturbations observed by a single ground-based magnetometer and actual current systems.

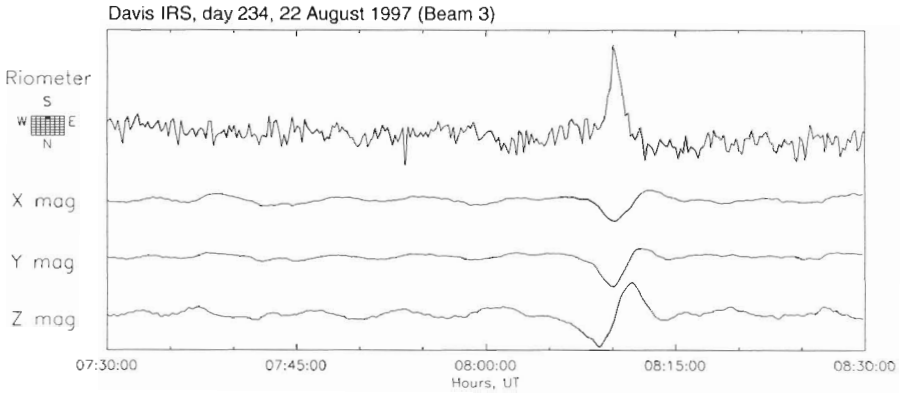


Figure 7. *Time series of received signal intensity in one beam of the SHIRE array, and the associated magnetic signature as measured by a co-located three-component fluxgate magnetometer on 22 August 1997. An absorption signature of comparable period can be seen to be associated with the magnetic impulse commencing at 0808 UT. A positive deflection in the riometer trace denotes an increase in radiowave absorption.*

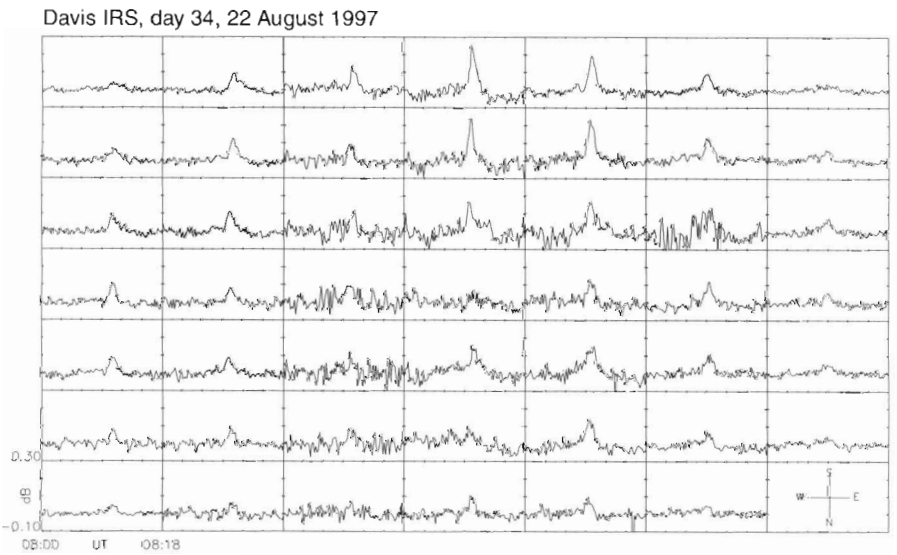


Figure 8. *Absorption intensities (in dB) in all beams of the SHIRE array during the interval 0800–0818 UT on 22 August 1997. The frames are arranged in the order that the beams project onto the ionosphere as viewed from below. Magnetic south is to the top of the page. A positive deflection in the riometer trace denotes an increase in radiowave absorption.*

The imaging riometer response here is representative of the more commonly observed absorption events associated with magnetic impulse events. Enhanced absorption is seen in Figure 8 to occur in all beams of the array over the duration of the event which commenced at 0808 UT (~ 1000 MLT) and finished at 0812 UT. Unlike the previous example, this event has a temporally unipolar riometer signature of comparable period to the magnetic impulse. The CNA patch can be seen in Figure 9 to form close to the eastern boundary of the instrument field-of-view and expand into the field-of-view over a time scale of ~ 1 minute. The patch is then seen to move west and south out of the field-of-view of the imaging riometer. Total absorption event duration is just under five minutes. Using the same cross-correlation technique described above, the phase velocity was determined to be ~ 2 km s^{-1} in the geomagnetic westward direction, and ~ 1 km s^{-1} in the southward (i.e. antisunward) direction.

Magnetic activity during the event was relatively quiet ($K_p = 1^+$), and again the inferred solar wind and IMF conditions at the magnetopause

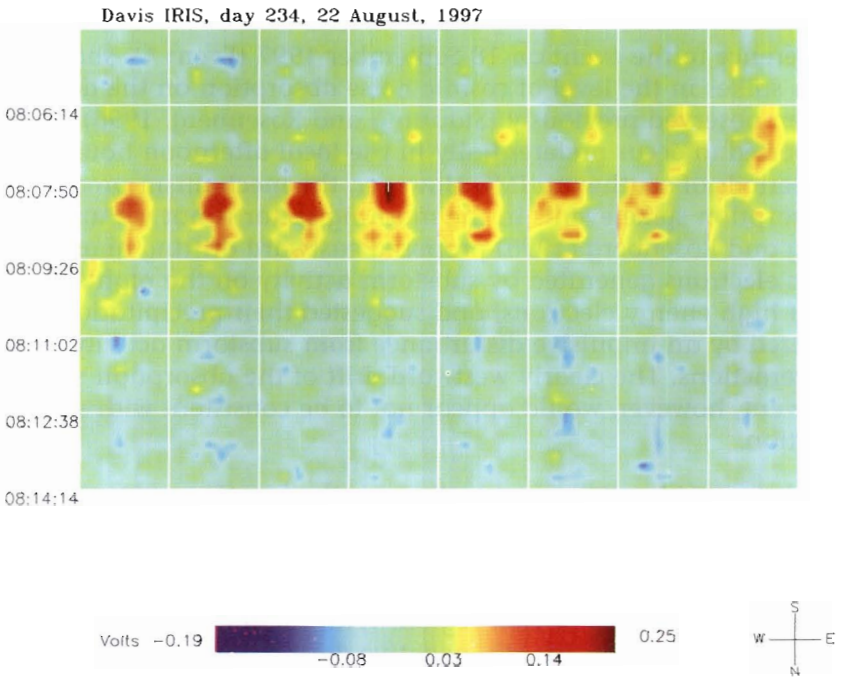


Figure 9. *Time series of absorption images (L to R, T to B) for the interval 0805–0813 UT on 22 August 1997 showing the progression of the MIE associated absorption event. Each image represents a ~ 200 x 200 km^2 view of the ionosphere.*

just prior to the event were stable. However, this time IMF B_z was marginally northward (+1 nT), and IMF B_y slightly positive (+2 nT). Under these conditions, we can reasonably expect Davis to be situated equatorward of the cusp and cleft on closed magnetic field lines. This is confirmed by an inspection of the dynamic power and phase spectra for this day which reveals the presence of strong and consistent Pc5 field-line resonance signatures in the magnetic field at Davis, as observed by a co-located induction magnetometer. For magnetic field-line resonance signatures the appearance of a 'Pc5 arch' in cross-power and the shift in relative phase for Pc5 pulsations between longitudinal stations about noon (Ables *et al.*, 1998) confirm the supposition that Davis is mapped to closed magnetic field lines on this day.

Although there is no evidence for a solar wind dynamic pressure variation impinging on the magnetosphere at the start of this event, the bipolar deflection of B_z, the large spatial scale of the absorption event and its antisunward motion on closed magnetic field lines at ~ 2 km s^{-1}, are all consistent with the signature of a Travelling Convection Vortex (TCV) as reported by Friis-Christensen *et al.* (1988).

11.4 Discussion

With reference to the event on 18 September 1997 (Event 1), short duration spikes in the level of cosmic noise absorption on the dayside have been observed previously (Stauning and Rosenberg, 1996). They were reported to occur preferentially in the local afternoon hours and during periods of southward IMF. However, the mechanism producing such absorption spike events in the morning sector is not yet understood. Stauning and Rosenberg (1996) proposed eastward drifting quasi-trapped energetic electrons generated by substorm activity on the nightside as the source of high energy electrons, and suggested their precipitation may be triggered by an impulsive disturbance from substorm activity or solar wind interactions. The clearly westward drift of the absorption signature in this event, however, would appear not to be consistent with such an explanation.

One further consideration of importance is the location of Davis with respect to the open/closed field line boundary. As the drift motion of the trapped population of electrons requires closed field lines, absorption events of the electron precipitation type should be restricted to regions equatorward of the cusp or cleft region. With the prevailing IMF B_z and B_y strongly and consistently negative during and immediately prior to the event on 18 September 1997, the footprint of the cusp would be expected to be shifted well into the post-noon sector away from Davis, and the polar cap expanded equatorward of Davis, leaving the station mapped to open field lines (Newell *et al.*, 1989). Confirming the supposition of open magnetic field lines is the absence of distinct Pc5 resonance signatures (such as the 'Pc5 arch') in the dynamic power spectrum for this day. Although the absence of such signatures is in itself no way conclusive evidence that the station was in the open field-line

regime, the combination of favourable IMF conditions and an absence of resonance signatures would suggest that this was indeed the case. In light of this, it would appear unlikely that eastward drifting populations of trapped substorm electrons were the source of energetic electrons producing the observed riometer event. Finally, WIND spacecraft measurements reveal no evidence of solar wind dynamic pressure variations impinging on the magnetopause prior to the event. However, it must be noted the WIND spacecraft was located 88 R_E upstream in the solar wind.

An interesting feature of the first event was the bipolar structure in the riometer signature. A depression in absorption levels such as that evident in this event, is best explained in terms of a modulation of radiowave absorption about a previously enhanced background level. Considering the two most likely mechanisms mediating the level of cosmic noise absorption, both electron precipitation and collisional heating of the lower E- region ionosphere could conceivably produce a bipolar fluctuation in CNA. However, earlier studies of CNA enhancements thought to be caused by collisional heating of the E-region (Stauning, 1984) have shown the characteristic absorption signature to be weak and slowly fluctuating, somewhat different from the rapidly changing spike event observed here.

With reference to the event on 22 August 1997 (Event 2), the appearance of distinct Pc5 resonance signatures concurrent with the event, confirms the event occurrence on closed magnetic field lines. Consistent antisunward motion of the plasma patch, then on closed magnetic field lines, suggests an explanation in terms of viscous interaction of magnetic flux tubes with solar-wind plasma flow down the flanks of the magnetosphere.

11.5 Summary

The recently installed SHIRE imaging riometer is being used to investigate the spatio-temporal morphology of cosmic noise absorption (CNA) events associated with magnetic impulse events at high-latitudes. Such observations enable further characterisation of magnetic impulses according to the dynamics and spatial scale of their associated absorption signatures. In this preliminary study, two distinct morphologies have been observed on the dayside:

(i) Localised, short duration absorption spikes restricted to a small number of beams, and often showing a bipolar structure. Possibly the ionospheric signature of an Flux Transfer Event (FTE), and

(ii) a widespread increase in absorption of spatial scale larger than the 200 x 200 km^2 field-of-view of the imaging riometer, consistent with the observed characteristics of Travelling Convection Vortices (TCV's).

Further study will involve a more precise identification of the location of the cusp and cleft footprints and the open/closed field line boundary in relation to Davis during each event. Use of DMSP particle data and a study of local magnetic pulsation activity during and prior to each event will allow reasonable estimates of the local magnetic field line topology, so that more accurate classification of events may be made.

11.6 Acknowledgments

This research was supported by the Australian Research Council, the University of Newcastle, the University of Maryland, the Australian Antarctic Science Advisory Committee, and the Australian Antarctic Division. We thank Lloyd Symons, and other ANARE staff for their efforts in installing SHIRE at Davis. The Cooperative Research Centre for Satellite Systems (CRCSS) is supported by the Commonwealth of Australia.

References

Ables, S.T., Fraser, B.J., Waters, C.L., Neudegg, D.A. and Morris, R.J. (1998). Monitoring cusp/cleft topology using Pc5 ULF waves. *Geophysical Research Letters,* 25: 1507–1510.

Butler, J. and Lowe, R. (1961). Beam-forming matrix simplifies design of electronically scanned antennas. *Electronic Design,* 12. p. 170.

Chaston, G.C., Hansen, H.J., Menk, F.W., Fraser. B.J. and Hu, Y.D. (1993). Ground signatures of convecting reconnected flux tubes. *Journal of Geophysical Research,* 98 (A11): 19151–19161.

Detrick, D.L. and Rosenberg, T.J. (1990). A phased-array radiowave imager for studies of cosmic noise absorption. *Radio Science,* 25 (4): 325–338.

Friis-Christensen, E., McHenry, M.A., Clauer, C.R. and Vennerstrom, S. (1988). Ionospheric traveling convection vortices observed near the polar cleft: a triggered response to sudden changes in the solar wind. *Geophysical Research Letters,* 15: 253–256.

Glassmeier, K.-H., Lester, M., Mier-Jedrzejovicz, W.A.C., Green, C.A., Rostoker, G., Orr, D., Wedeken, U., Junginger, H. and Amata, E. (1984). Pc5 pulsations and their possible source mechanisms: a case study. *Journal of Geophysics,* 55: 108–119.

Glassmeier, K.-H. (1992). Traveling magnetospheric convection twin-vortices: observations and theory. *Annales Geophysicae,* 10: 547–565.

Heikkila, W.J., Jorgensen, T.S., Lanzerotti, L.J. and Maclennan, C.G. (1989). A transient auroral event on the dayside. *Journal of Geophysical Research,* 94: 15291–15305.

Korotova, G.I., Rosenberg, T.J., Lanzerotti, L.J. and Weatherwax, A.T. (1999). Cosmic noise absorption at South Pole station during magnetic impulse events. *Journal of Geophysical Research,* 104 (A5): 10327–10334.

Lanzerotti, L.J., Lee, L.C., Maclennan, C.G., Medford, L.V., Lee, L.C. and Wolfe, A. (1986). Possible evidence of flux transfer events in the polar ionosphere. *Geophysical Research Letters,* 13: 1089–1092.

Lanzerotti, L.J., Maclennan, C.G., Konik, R.M., Wolfe, A. and Venkatesan, D. (1991). Cusp latitude magnetic impulse events 1. Occurrence statistics. *Journal of Geophysical Research,* 96(A8): 14009–14022.

McHenry, M. and Clauer, C.R. (1987). Modeled ground magnetic signatures of flux transfer events. *Journal of Geophysical Research,* 92: 11231–11240.

McHenry, M.A., Clauer, C.R. and Friis-Christensen, E. (1990). Relationship of solar wind parameters to continuous dayside, high-latitude traveling ionospheric convection vortices. *Journal of Geophysical Research,* 95: 15007–15022.

Newell, P.T., Meng, C-I., Sibeck, D.G. and Lepping, R. (1989). Some low-altitude cusp dependencies on the interplanetary magnetic field. *Journal of Geophysical Research,* 94(A7): 8921–8927.

Russell, C.T. and Elphic, R.C. (1979). ISEE observations of flux transfer events at the dayside magnetopause. *Geophysical Research Letters,* 6: 33–36.

Sibeck, D.G., Baumjohann, W. and Lopez, R.E. (1989). Solar wind dynamic pressure variations and transient magnetospheric signals. *Geophysical Research Letters,* 16: 13–16.

Southwood, D.J. (1987). The ionospheric signature of flux transfer events. *Journal of Geophysical Research,* 92: 3207–3213.

Stauning, P. (1984). Absorption of cosmic noise in the E-region during electron heating events: a new class of riometer absorption events. *Geophysical Research Letters,* 11: 1184–1187.

Stauning, P. (1985). In: Holtet, J.A. and Egeland, A. (Eds.). *The Polar Cusp.* Pp. 365–376.

Stauning, P., Clauer, C.R., Rosenberg, T.J., Friis-Christensen, E. and Sitar, R. (1995). Observations of solar-wind-driven progression of interplanetary magnetic field By-related dayside ionospheric disturbances. *Journal of Geophysical Research,* 100(A5): 7567–7568.

Stauning, P. and Rosenberg, T.J. (1996). High-latitude daytime absorption spike events. *Journal of Geophysical Research,* 101(A2): 2377–2396.

Zhu, L., Gifford, P., Sojka, J.J. and Schunk, R.W. (1997). Model study of ground magnetic signatures of traveling convection vortices. *Journal of Geophysical Research,* 102(A4): 7449–7459.

Zhu, L., Schunk, R.W. and Sojka, J.J. (1999). Effects of magnetospheric precipitation and ionospheric conductivity on the ground magnetic signatures of traveling convection vortices. *Journal of Geophysical Research,* 104(A4): 6773–6782.

12. DYNAMICS OF THE POLAR CAP IONOSPHERE PART 1. APPLICATIONS OF DIGISONDE ANIMATIONS IN IONOSPHERIC RESEARCH

A.M. Breed[1], T.M. Maddern[1], P.L. Dyson[2] and R.J. Morris[1]

[1] Atmospheric and Space Physics
Australian Antarctic Division
Kingston Tasmania 7050 Australia
(e-mail: anthony.breed@aad.gov.au)

[2] Department of Physics
La Trobe University
Bundoora Victoria 3083 Australia

Abstract

A UMLCAR (University of Massachusetts Lowell, Centre for Atmospheric Research) Digital Ionosonde (DPS-4) has operated at the Australian Antarctic polar cap station Casey (80.8°S geomagnetic latitude) since early 1993, primarily to study the dynamics of the southern polar cap ionosphere. To gain a better understanding of the ionospheric motions involved and to identify dynamical features, animated data displays have been developed using ionospheric drift and ionogram data combined together with other geophysical data sets. These displays generally show ionospheric skymaps (echo locations), drift velocities and Doppler ionograms (vertical profiles), produced with a time resolution of up to a frame every three minutes. Static displays of other geophysical data sets (e.g. Total Electron Content (TEC) from Global Positioning System (GPS) satellites or Interplanetary Magnetic Field (IMF) measurements) can be incorporated into the animations. Ionospheric changes can then be viewed frame-by-frame using commercial animation software.

The displays provide a simple means of displaying large amounts of data and identifying, for example, drifting polar patches, or the response of ionospheric drift to changes in the IMF. Since the ionosonde takes measurements throughout the bottomside ionosphere, different height regions (e.g. E and F) can be examined independently or compared directly. Doppler and directional information featured in the animations provide a powerful technique for studying the evolution of ionospheric phenomena such as polar patches.

Some of the applications of the 'ionomovie' animations in ionospheric research will be demonstrated.

12.1 Introduction

The Digisonde Portable Sounder (DPS-4) (Reinisch *et al.*, 1995), like most modern digital ionosondes, is able to provide high temporal and spatial resolution measurements of the ionosphere. These measurements contain a wealth of information, including, Doppler shifts, angles-of-arrival and group ranges, from significant numbers of ionospheric echo sources simultaneously. The data sets thus generated can be of considerable size, requiring complex and time consuming computer analysis. Initial display techniques essentially produce 'one-dimensional' drift velocity or azimuth versus time plots (as in Buchau *et al.*, 1988). These combined measurements over all heights and frequencies produce average values at a given time. Recently, significant advances have been made in the development of new display methods of digisonde data for ionospheric research (Parkinson *et al.*, 1999a), making good use of colour as a means of identifying specific features in the plots, and adding a range/height dimension. However, these types of visualisation packages are still limited to generating static displays whose complexity and content can make it difficult to study smaller scale or transient features present in the data.

This paper describes a method to provide an alternative view of digisonde data by using animation techniques to better demonstrate the dynamic nature of the ionosphere and ionospheric features. By incorporating drift data and ionograms, and displaying multiple ranges, the new displays can give a virtual three-dimensional view of the ionosphere, and how it changes with time.

The use of animations in ionospheric research is not new, with early ionosondes recording ionogram images on film that could then be projected. Computers have increased the flexibility of this technique, but still display a series of ionograms only, an approach that provides a limited view of the information available (giving effectively only a simple vertical profile). The animation methods described here (tentatively called 'ionomovies') add a horizontal dimension to the animation by incorporating 'skymap' displays (showing spatial distribution and Doppler shift of ionospheric echo sources in a horizontal view), generated by the DPS-4 Drift Data Analysis (DDA) software (Scali *et al.*, 1995). The DDA software locates echo sources by applying a Doppler sorted interferometry (DSI) technique to measurements taken during a 'drift mode' sounding program of the DPS-4. Echo sources may be sorted according to vertical group range (disregarding refraction effects), allowing skymaps to be generated for specific height ranges (e.g. E- and F-regions). Corresponding velocity vectors deduced from the skymaps may also be included on the displays.

The DPS-4 has several ionogram modes with the most commonly used being beamforming (BEM) ionograms storing coarse angle-of-arrival and single level Doppler information. Concurrent ionograms and drift files may be included in the same 'frame' of the ionomovies. With the current configuration of the DPS-4, a cycle of one ionogram followed by a drift

measurement can be completed every three minutes. Higher repetition rates may be available with future upgrades.

To increase the usefulness of the ionomovies in investigating ionospheric features and their evolution, they were designed to allow simple comparison with other geophysical parameters. By making the animation displays flexible enough to allow the inclusion of other data sets (e.g. TEC, satellite scintillation or IMF), the interrelationships between these parameters and dynamic features in the ionospheric environment can be investigated.

Although applicable to any research of dynamic or transient events in the ionosphere, this study demonstrates the use of ionomovies in the identification and/or characterisation of sporadic-E events or polar patches.

12.2 DPS operation and analysis

The standard operating mode of the Casey DPS-4 involves running an ionogram (usually beamforming) program every ten minutes, followed by a number of drift programs covering a range of frequencies and heights. For best temporal resolution of features using the ionomovies, an operating mode was devised that completed an ionogram and a basic drift program every three minutes. It is expected that, with future upgrades of the Casey DPS, this time may be reduced to a two or even one minute cycle.

BEM ionograms are recorded by using the four receive antennas to form seven 'beams' into which the strongest ionospheric echoes (at each selected frequency and range) are sorted. One beam contains vertical echoes (which are further separated into ordinary and extraordinary modes) out to a zenith angle of 20 degrees, and the six remaining beams include off vertical reflections in 60 degree azimuthal segments. The 'beam number' is stored in the data file along with amplitude and coarse Doppler (i.e. positive or negative) information.

Drift mode measurements record up to 128 complex samples (depending on desired Doppler frequency resolution) from each receive antenna for each frequency and range bin. These undergo a Fourier transform to obtain Doppler and phase spectra, and are subsequently stored in a data file. With further processing using the DDA package (Scali *et al.*, 1995a), echo sources are obtained from Doppler Sorted Interferometry (DSI) (Scali *et al.*, 1995b) whereby echoes are sorted according to Doppler shift, and directions of arrival are derived from phase measurements over the four antennae of the receive array. This results in a SKY data file containing echo coordinates, ranges and Doppler shifts from which Skymaps are produced. Further processing with DDA on these SKY files yields 'drift velocities' deduced by a least-squares fit to at least three echoes.

This procedure for determining drift velocities depends on a number of assumptions. It assumes minimal refraction effects. It also assumes that the ionosphere is moving with a uniform, bulk flow, and that the echo sources detected are imbedded in that bulk flow. The validity and limitations of digisonde drift measurements is a subject of some debate. Comparative studies at high latitudes have shown that digital ionosondes basically measure the ion motion in the F-region (Grant *et al.*, 1995; Smith *et al.*, 1998). However, Monselesan and Morris (1999), discuss possible discrepancies in drift measurements introduced by the presence of plasma waves or irregularities, primarily in the E-region, but also of importance at F-region heights. It is noted that the current study does not attempt to address these possible limitations, but simply, to show a new method of visualising digisonde data from which other conclusions can then be drawn.

The DPS data were converted to animated displays via a number of steps using software packages developed for the purpose. Firstly, the DPS binary files (ionograms and SKY files) are converted to a text format to enable easier sorting and display. At this point, approximate group heights of echoes from the SKY files are calculated using group range and elevation values contained in the data files. Using data from a SKY file, a plot is created displaying Skymaps (horizontal locations) of ionospheric echoes sorted into two virtual height ranges (on separate axes) and colour coded according to Doppler shift. The corresponding ionogram (recorded immediately before the drift file) is filtered by amplitude and colour coded according to beam direction and polarisation, and displayed in a separate panel on the plot. Any other data that may be relevant to the feature being studied (e.g. IMF or TEC) are then displayed in a further separate panel on the same plot. The resulting image is saved as a GIF (Graphical Interchange Format) image file and merged with all subsequent images using a commercial animation package to form the final display.

Figure 1 shows an example of one such image from an animation sequence. The left top panel in the figure shows echoes in a horizontal plane in the virtual height range 90 to 200 km, approximately E-region. The right top panel shows the range 200 to 600 km (F-region). Echoes are shown as either cyan (positive Doppler) or red (negative Doppler). The outer circle of the plots represents a zenith angle of 45 degrees. The approximate solar azimuth is shown as a star symbol on the outer circle. Superimposed on each Skymap plot is the calculated drift velocity vector for that height range (the outer circle represents a velocity of 1500 m s^{-1}).

The lower left panel shows the corresponding ionogram, while the lower right panel, in this case, shows the three orthogonal interplanetary magnetic field values (in Geocentric Solar-Magnetospheric (GSM) coordinates). All directions shown are in magnetic compass coordinates (approximately -92 degrees deviation from geographic directions).

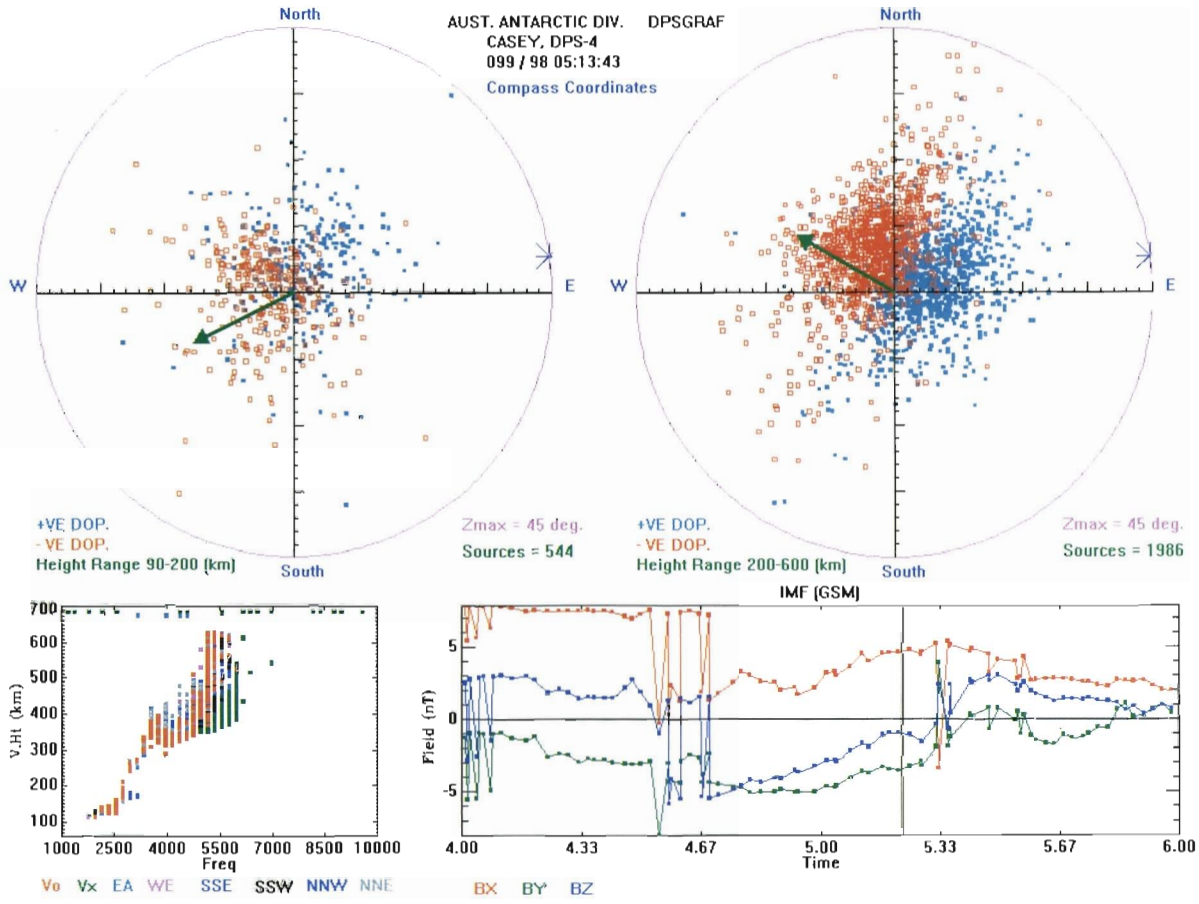


Figure 1. *Example of one frame of an ionomovie, 9 April 1998.*

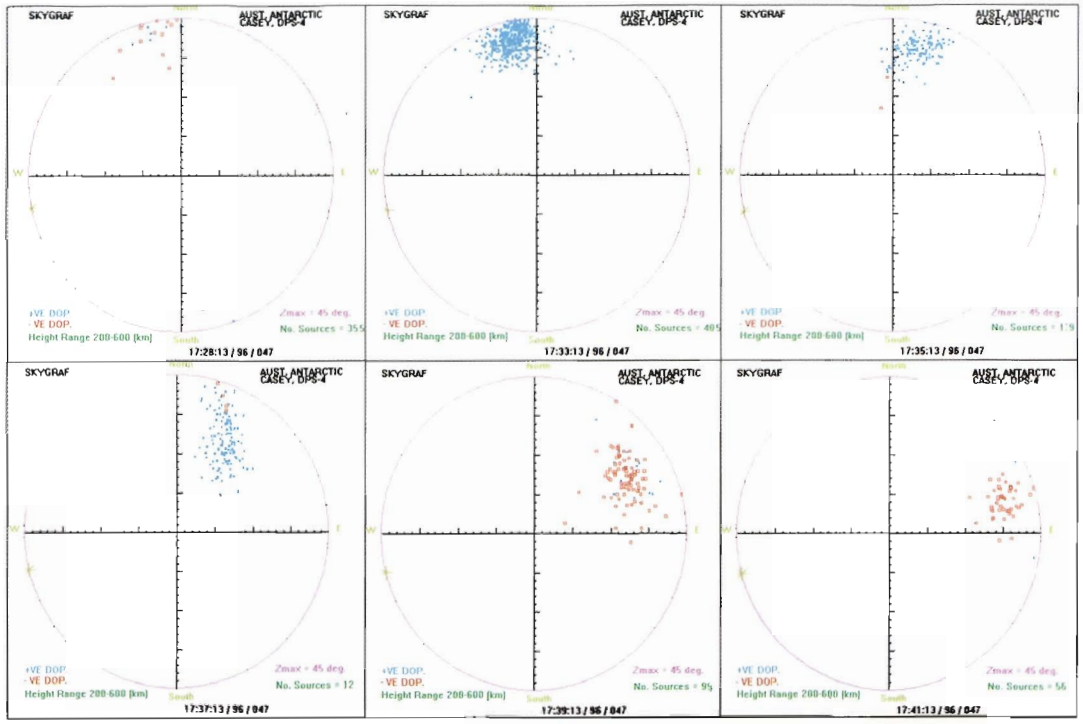


Figure 2. *Passage of a small polar patch through the field-of-view of the DPS-4 at Casey, 16 February 1996.*

12.3 Applications and examples

The main idea behind the development of the ionomovies, apart from simplifying the display and analysis of large volumes of data, was to assist in a study on polar patches. Polar patches are regions of increased plasma density, up to 1000 km in horizontal extent, which convect across the polar cap in an antisunward direction. The ionosonde signatures of a patch are significant increases in critical frequencies on ionograms or U-shaped traces on constant frequency, range versus time plots (James and MacDougall, 1997; MacDougall *et al.*, 1996). It was expected that the ionomovie technique could also be used to identify and further investigate the formation and motion of polar patches. Although results from this study are still very preliminary, a number of patches have been identified in ionomovie animations. Figure 2 shows an example of a small patch convecting through the field-of-view of the DPS. The figure shows just the Skymap image sequence from an ionomovie, clearly showing the passage of a patch to the north-east (compass) of the station. The Doppler shift is seen to change sign as the patch passes the point of closest approach. This patch was also identified as a U-shaped trace in the data of Parkinson *et al.* (1999b). Larger patches tend to be less easily identified than this example and appear to take the form of 'fronts' (the leading or trailing edges of the patch) passing across the field-of-view of the instrument.

Other possible applications of the ionomovie technique could include a detailed investigation of drift velocity shear in the ionosphere (by the choice of appropriate virtual height ranges) and the possible association with the formation of sporadic-E layers (Parkinson *et al.*, 1997; Parkinson *et al.*, 1998). Examples have been seen where the azimuths of the drift vectors in the approximate E- and F-region Skymaps have shown up to 180 degrees difference.

12.4 Conclusions

A new dynamic display technique of Digisonde DPS-4 data has been developed in the form of ionomovie animations. This has great potential for studying the formation of dynamics of small or transient features and motions in the ionosphere. In this paper, an example of how this technique can help in the identification and study of polar patches has been presented. It is anticipated that, with the flexibility allowed in this technique, it will prove useful in many other areas of ionospheric study including sporadic-E and ionospheric drift-shear, convection pattern changes with IMF, and ionospheric irregularities and instabilities associated with the high-latitude ionosphere.

References

Buchau, J., Reinisch, B.W., Anderson, D.N., Weber, W.J. and Dozois, C. (1988). Polar cap plasma convection measurements and their relevance to the modelling of the high-latitude ionosphere. *Radio Science*, 23: 521–536.

Grant, I.F., MacDougall, J.W., Ruohoniemi, J.M., Bristow, W.A., Sofko, G.J., Koehler, J.A., Danskin, D. and Andre, D. (1995). Comparison of plasma flow velocities determined by the ionosonde Doppler drift technique, SuperDARN radars and patch motion. *Radio Science,* 30: 1537–1549.

James, H.G. and MacDougall, J.W. (1997). Signatures of polar cap patches in ground ionosonde data. *Radio Science,* 32: 497–513.

MacDougall, J.W., Grant, I.F. and Hamza, A. (1996). Velocity fluctuations associated with polar patches. *Radio Science,* 31: 595–606.

Moselesan, D.P. and Morris, R.J. (2001). Digital ionosonde observation of E/F-regions during intense lacuna conditions at polar cap latitude: implications for drift velocity determination. (in preparation).

Parkinson, M.L, Monselesan, D.P., Smith, P.R., Dyson, P.L. and Morris, R.J. (1997). Digital ionosonde measurements of the height variation of drift velocity in the southern polar cap ionosphere: Initial results. *Journal of Geophysical Research,* 102: 24075–24090.

Parkinson, M.L., Dyson, P.L., Monselesan, D.P. and Morris, R.J. (1998). On the role of electric field direction in the formation of sporadic E-layers in the southern polar cap ionosphere. *Journal of Atmospheric and Solar Terrestrial Physics,* 60: 471–491.

Parkinson, M.L., Breed, A.M., Dyson, P.L. and Morris, R.J. (1999a). Application of the Dopplionogram to Doppler-sorted interferometry measurements of ionospheric velocity. *Radio Science,* 34: 899–912.

Parkinson, M.L., Monselesan, D.P., Breed, A.M., Dyson, P.L. and Morris, R.J. (1999b). Signatures of the ionospheric cusp in digital ionosonde measurements of plasma drift above Casey, Antarctica. *Journal of Geophysical Research,* 104: 22487–22498.

Reinisch, B.W., Bullett, T.W., Scali, J.L. and Haines, D.M. (1995). High latitude digisonde measurements and their relevance to IRI. *Advances in Space Research,* 16: 17–26.

Scali, J., Reinisch, B., Dozois, C., Bibl, K., Kitrosser, D., Haines, M. and Bullett, T. (1995a). *Digisonde Drift Analysis.* University of Massachusetts, Lowell Center for Atmospheric Research.

Scali, J., Reinisch, B.W., Heinselman, C.J. and Bullett, T.W. (1995b). Coordinated digisonde and incoherent scatter radar F-region drift measurements at Sondre Stromfjord. *Radio Science,* 30: 1481–1498.

Smith, P.R., Dyson, P.L., Monselesan, D.P. and Morris, R.J. (1998). Ionospheric convection at Casey, a southern polar cap station. *Journal of Geophysical Research,* 103: 2209–2218.

13. DYNAMICS OF THE POLAR CAP IONOSPHERE PART2. CASE STUDIES OF POLAR PATCHES ABOVE CASEY, ANTARCTICA

A.M. Breed[1], T.M. Maddern[1], P.L. Dyson[2] and R.J Morris[1]

[1] Atmospheric and Space Physics
Australian Antarctic Division
Kingston Tasmania 7050 Australia
(email: anthony.breed@aad.gov.au)

[2] Department of Physics
La Trobe University
Bundoora Victoria 3083 Australia

Abstract

During 1997 and early 1998 a series of campaigns were undertaken at the Australian Antarctic polar cap station Casey (-80.8° geomagnetic latitude) with the aim of identifying and gaining further understanding of polar patches. The main instrument used for these campaigns was a University of Massachusetts Lowell, Centre for Atmospheric Research (UMLCAR) Digital Ionosonde (DPS-4), running a three-minute cycle of Doppler ionograms and drift velocity measurements. Patches identified from group range and critical frequencies on the ionogram records were compared with drift measurements and other geophysical data sets. Good correlations were observed with Total Electron Content (TEC) derived from Global Positioning System (GPS) satellite observations with a receiver also located at Casey. The field-of-view for GPS observations allowed patches to be observed over large distances, providing some insight into their origin. The influence of the Interplanetary Magnetic Field (IMF) on the formation and occurrence of patches was also considered. Comparison with plasma drift velocities, (derived from DPS drift measurements), confirmed that patches are associated with fluctuations in horizontal drift velocity. In the current study, peaks in the magnitude of the horizontal drift velocity correlate with the patch edges. This paper considers case studies of three days of patch observations in April 1998.

13.1 Introduction

13.1.1 Morphology

The term polar patch (hereafter referred to as patches) is used to refer to enhanced regions of ionospheric F-layer plasma density, convecting across the polar cap in a generally antisunward direction. (Similar enhancements observed outside the polar cap are generally called auroral blobs.) The patches have horizontal scale sizes of 100–1000 km and are observed travelling at velocities in the range 300–1000 m s^{-1}.

The patches themselves often contain smaller scale structure and irregularities (1–10 km range), particularly on the edges as determined from scintillation effects.

The first observations of these phenomena were made by Hill (1963), but it appears no intensive studies of patches were undertaken until those reported by Buchau *et al.* (1983, 1985) and Weber *et al.* (1984, 1986). In-depth reviews of polar patch morphology, dynamics and formation processes are given by Crowley (1996) and Rodger (1998).

Patches are observed to have plasma densities comparable with the dayside sub-auroral ionosphere, suggesting that they originate in this region and are subsequently transported through the cusp region to the polar cap by the high-latitude convection. Studies have shown that southward interplanetary magnetic field (IMF) (B_Z negative) is conducive to the formation of patches, as is a disturbed geomagnetic field (Kp > 4) (Buchau *et al.*, 1983; Weber *et al.*, 1984). Patches may also form simultaneously in geomagnetically conjugate regions (Rodger et al., 1994b). The electron temperature of patches (determined from incoherent scatter radar measurements) is low (Anderson *et al.*, 1988) indicating that they are not caused by particle precipitation.

Diurnal and seasonal occurrence rates of patches are discussed by Rodger and Graham (1996). Their data indicated maxima in occurrence around magnetic noon between equinox and winter seasons.

13.1.2 Formation mechanisms

A number of possible formation mechanisms of polar patches have been put forward and discussed or modelled in the literature. Although there is no universal agreement on the validity of the individual mechanisms (some depend on how a polar patch is specifically defined), there does appear to be general agreement that the source of ionisation for patch formation is the enhanced ridge of plasma called the 'tongue of ionisation' created by the polar convection dragging sunlit plasma equatorward of the cusp into the polar cap (Steele and Cogger, 1996). The method by which this plasma is 'broken up' into patches is a source of debate, although overall there seems to be approximately three main themes. They are :

a. The rapid expansion or change in the polar convection patterns associated with either IMF B_Z transitions (Valladares *et al.*, 1998) or transient magnetopause reconnection (Flux Transfer Events (FTE's)) (Lockwood and Carlson, 1992), 'modulating' the plasma flow from the dayside into patches.

b. The distortion and fragmentation of an existing steady plasma ridge (tongue of ionisation) by rapid changes in IMF B_Y and possibly gravity waves (Sojka *et al.*, 1993).

c. The 'chopping' of an existing, steady plasma ridge into patches by fast plasma jets (Flow Channel Events (FCE's)) also associated with FTE's. The increased plasma velocities (several kilometres per

second) decrease the expected F-region plasma lifetime, enhancing recombination and causing depletions in the plasma ridge (Rodger *et al.*, 1994a; Valladares *et al.*, 1994, 1996). This appears to be the most widely accepted formation mechanism.

13.1.3 Signatures and detection

Polar patches are detected by a wide range of ground-based instruments and have a correspondingly wide array of signatures. Most early observations relied on an enhanced 630 nm optical emission detected with photometers or imaged with intensified all-sky cameras. A number of recent studies have still used this technique of detection (Fukui *et al.*, 1994; McEwen *et al.*, 1995a, 1995b). Sojka *et al.* (1997) discuss problems comparing optical measurements with radio frequency observations.

More recently, signatures of patches have been detected with satellite-borne instruments either by *in situ* sampling (Coley and Heelis, 1995; Kivanc and Heelis, 1997) or through remote sensing using visible, ultraviolet or X-ray imagers (e.g. on board the POLAR satellite). Also involving satellites are the techniques of Total Electron Content (TEC) measurements, using Global Positioning System (GPS) or Navy Navigation Satellite System (NNSS) satellite beacons, which allow patches to be detected at ground receivers as large increases in line-of-sight electron content or radio scintillation on satellite signals (Weber *et al.*, 1986).

Ground-based radio frequency signatures include increased absorption measurements on riometers (Rosenberg *et al.*, 1993; Wang *et al.*, 1994) and electron density structures observed with HF radars (incoherent scatter radar or coherent scatter radar) (Ogawa *et al.*, 1998a, 1998b).

Of more particular importance to the current studies are the signatures of patches observed on digital ionosondes. The most obvious effect of polar patches seen on ionosonde measurements are increases in critical frequency (f_0F_2) associated with the increased plasma density. However, a more definitive signature appears in plots of h't, (virtual height versus time at a constant frequency). James and MacDougall (1997) have modelled the effects of electron density structures on ionosonde transmissions and predict the appearance of 'U-shaped' (or hyperbolic) traces in ionograms and h't plots. These U-shaped traces are shown in the studies of MacDougall *et al.* (1996), where they are the primary method of patch identification. The formation of the U-shaped trace relies partly on the fact that echoes observed with the digisonde are not necessarily from overhead. As a high-density patch approaches the sounder, echoes are picked up at large, oblique ranges. The range of these echoes decreases as the patch approaches, and then increase again as it recedes from the station, hence producing the U-shape in the h't plots. Figure 1 of MacDougall *et al.* (1996), shows examples of the patch associated U-shaped plots from measurements made with a Canadian Advanced Digital Ionosonde (CADI). With digital ionosondes (such as the DPS-4) parameters such as direction of arrival and Doppler shift of

reflected signals are also available, and can further contribute to the identification of patches in data records.

In addition to these methods of detection, the DPS-4 also has the facility for drift mode operation, in which high resolution measurements of echo location (angle-of-arrival) and Doppler shift can be made over a number of echo sources, heights and frequencies, by the method of Doppler sorted interferometry (DSI) (Reinisch *et al.*, 1995; Scali *et al.*, 1995), effectively allowing 'Skymaps' of echo sources to be produced. Although the signature of patches in this type of display is still under investigation, it is expected that the edges of large patches would appear as 'fronts' of echoes crossing the field-of-view, associated with the presence of small

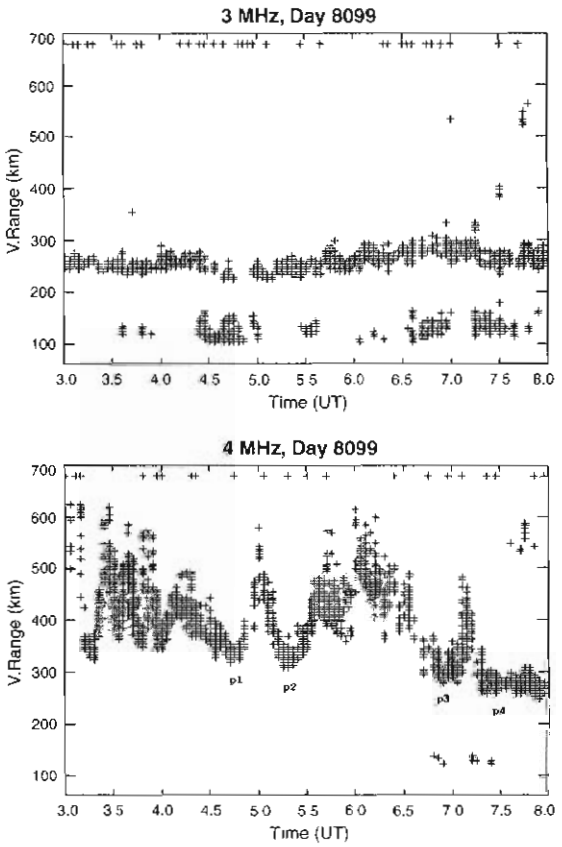


Figure 1. *A series of four patches identified on Casey Digisonde records for 9 April 1998. Virtual ranges are shown for both 3 and 4 MHz transmission frequencies.*

scale irregularities, while smaller patches may be seen in their entirety (Breed *et al.*, 1999).

13.1.4 Current studies

The current paper presents and discusses some case studies of polar patches, concentrating on patches identified on three days of a two-week campaign in April 1998. Some mention is also made of patches identified at other times. Patches were identified from the digisonde measurements visually by the U-shaped trace described above. Measurements of critical frequency, TEC, drift velocity and 630 nm optical emissions were then examined for corresponding patch signatures.

13.2 Instruments and analysis

The experiments described in this paper were located at Casey station, in the Australian Antarctic Territory. Casey lies at 66°17'S and 110°32'E geographic, (-80.8° geomagnetic latitude). Magnetic local time (MLT) at Casey is approximately UT + 5.5 hours.

The primary instrument used in this study is the UMLCAR Digisonde DPS-4. This instrument operates as a pulsed, Doppler radar, and consists of crossed rhombic transmit antennae and four crossed loop receive antennae (and associated independent receivers) in a triangular array, allowing measurements of angle-of-arrival, polarisation, and Doppler shifts of ionospheric echoes. For the patch campaign in April 1998, the DPS-4 was running an operating schedule consisting of a beam-forming ionogram (BEM) followed by a drift mode program, repeated every three minutes.

Beam-forming ionograms use the directional capabilities of the DPS-4 receive array to sort returning signals (maximum amplitude signal in each range and frequency bin) into one of seven directional 'beams' (consisting of one 'vertical' (to 20 degrees from zenith) beam and six off-vertical beams (between 20 and 45 degrees from zenith)). The vertical beam is further divided into ordinary and extraordinary polarisations. Coarse (one bit) Doppler information is also assigned to the echoes.

In DPS-4 drift mode sounding, echoes are sorted by Doppler shift (at a given frequency and virtual range), and accurate angles-of-arrival are obtained using phase shifts across the four receive antennae (hence the term Doppler sorted interferometry (DSI)). By varying the number of samples and the sample rate (by changing the number of pulses transmitted and the pulse repetition rate) the Doppler resolution and range of the Fourier transformed time series, and the number of possible echo sources (up to a maximum of 128 per range and frequency), can be controlled. Hence, from the drift mode measurements, a three dimensional 'Skymap' of echo sources and associated Doppler shifts above the sounder can be constructed. Using a least-squares-fit to at least three echoes, apparent 'drift' velocities are calculated. The analysis is carried out using the Digisonde Drift Analysis (DDA) software package (Reinisch *et al.*, 1995; Scali *et al.*, 1995; Parkinson *et al.*, 1999a).

To assist in identifying polar patches in digisonde data a suite of programs were written to analyse and display particular data parameters. This involved extracting and plotting critical frequencies (f_oF_2 and F_xI) and virtual ranges (at fixed frequencies) from the BEM ionograms, using colour to show Doppler signs or azimuths of echoes. Drift data were used directly to obtain ionospheric velocities or displayed in the animated 'ionomovie' format (Breed *et al.*, 2001).

Data from co-located GPS receivers were used to obtain TEC values and scintillation measurements. TEC measurements were corrected for satellite and receiver biases, and transformed from slant to vertical TEC values. This transformation assumed a median ionospheric height of 400 km to define a sub-ionospheric point and used the cosine of the zenith angle to project slant TEC to vertical TEC values (Breed *et al.*, 1998). The sub-ionospheric point is the spot on the Earth's surface directly below the intersection of the receiver-to-satellite path and 400 km altitude. Scintillation measurements were corrected according to the method of Nichols *et al.* (1999). Due to the inclination of the GPS satellite orbits, and an imposed zenith angle limit of 50 degrees, the majority of measurements with these receivers applied to sub-ionospheric points several degrees geographic north of Casey station.

Interplanetary magnetic field (IMF) component values, in Geocentric Solar Magnetic (GSM) coordinates, were obtained from WIND spacecraft data (1.5 minute averages). These were corrected to allow for a propagation time delay from the spacecraft to the Earth's bow-shock (assumed to be at 10 Earth radii) by using a simple approximation of distance divided by solar wind speed (also obtained from WIND spacecraft data). No allowance was made for any subsequent travel time caused by the slowing down of the solar wind close to the bow-shock or propagation time from the outer magnetosphere to the ionosphere which may contribute up to another 10 minutes to the dealy.

Simple optical measurements at 630 nm from a wide-angle photometer and ionospheric absorption measurements from a standard, 30 MHz riometer at the station were also available, and both data sets will be considered briefly.

Other data available at Casey station, but not used in the present study, are induction and fluxgate magnetometers, and an all-sky intensified video imager. These may be included in ongoing polar patch research and reported in future papers.

13.3 Case studies

The case studies discussed in this paper were from three days in April 1998. These intervals were not the only observations of polar patches in this period, but were picked as times displaying examples of large, well-defined patches, demonstrating their observation in DPS measurements and showing clear correlations with other data sets.

13.3.1 9 April 1998

Figure 1 shows the first case study of a series of patches observed on day 99 (9 April 1998). (Titles on figures of the form 'Day 8099' are derived from the digisonde data file system and consist of a single-character year and three-character day number, e.g. Year 1998, Day 099). Figure 1 shows the h't plots (group range as a function of time) for the frequencies 3 and 4 MHz, over a period of 5 hours. The patches are identified from U-shaped traces on the 4 MHz plot, as there is minimal effect on the 3 MHz group ranges in the current example. This arises because at this time (0400–0600 UT) on day 99 the 3 MHz signal was being reflected in the F1-layer (and on occasion the E-layer), and hence was not affected by the patches that appear to only disturb F2-layer heights (4 MHz signal). This is not always the case, and depends on time of day and the background (normal) electron densities. At night or times of low background density, the patches are more likely to be seen at lower frequencies. The interpretation of patches seen on 4 MHz range plots but not in 3 MHz plots, are given in Figure 1, we have indicated four distinct patches. However, the critical frequency f_0F_2 enhancements (see Figure 2) suggest that there may be more, less distinct patches present in the data.

Figure 2 shows the 4 MHz trace from Figure 1, combined with f_0F_2 values (also derived from DPS ionograms, with three point smoothing applied) and TEC derived from GPS satellite reception (averaged and restricted to satellites with elevation angles greater than 50 degrees). A comparison between these three data sets must take into account that they have different fields-of-view. The h't plot shows all 4 MHz echoes down to an angle of 45 degrees off vertical, while f_0F_2 measurements are restricted to angles less than 20 degrees. This should have little effect on the comparison between these plots except that patches may be seen first on the h't results. GPS TEC measurements relate to a small region 2–3 degrees geographic north of Casey, and hence, depending on time of day and direction of patch drift, may exhibit patch signatures some time before or after DPS measurements.

The dashed vertical lines on Figure 2 have been added to allow easy comparison between the three plots and indicate approximate times of closest approach to the station for the three patches. These times are estimated from the h't plots as the minima of the U-shaped features.

It is immediately obvious that the patch occurrences correspond to peaks in critical frequency. It is noted that there are a number of closely spaced peaks in f_0F_2 prior to 0400 UT, most likely indicating a series of smaller patches. Evidence of these is also present in the h't plot, although much less distinct than the later patches.

The lower plot indicates increases in TEC assumed to correspond to all three patches indicated. There are two or more distinct traces shown in this panel corresponding to different GPS satellites and hence different observation locations (although still within a relatively small region

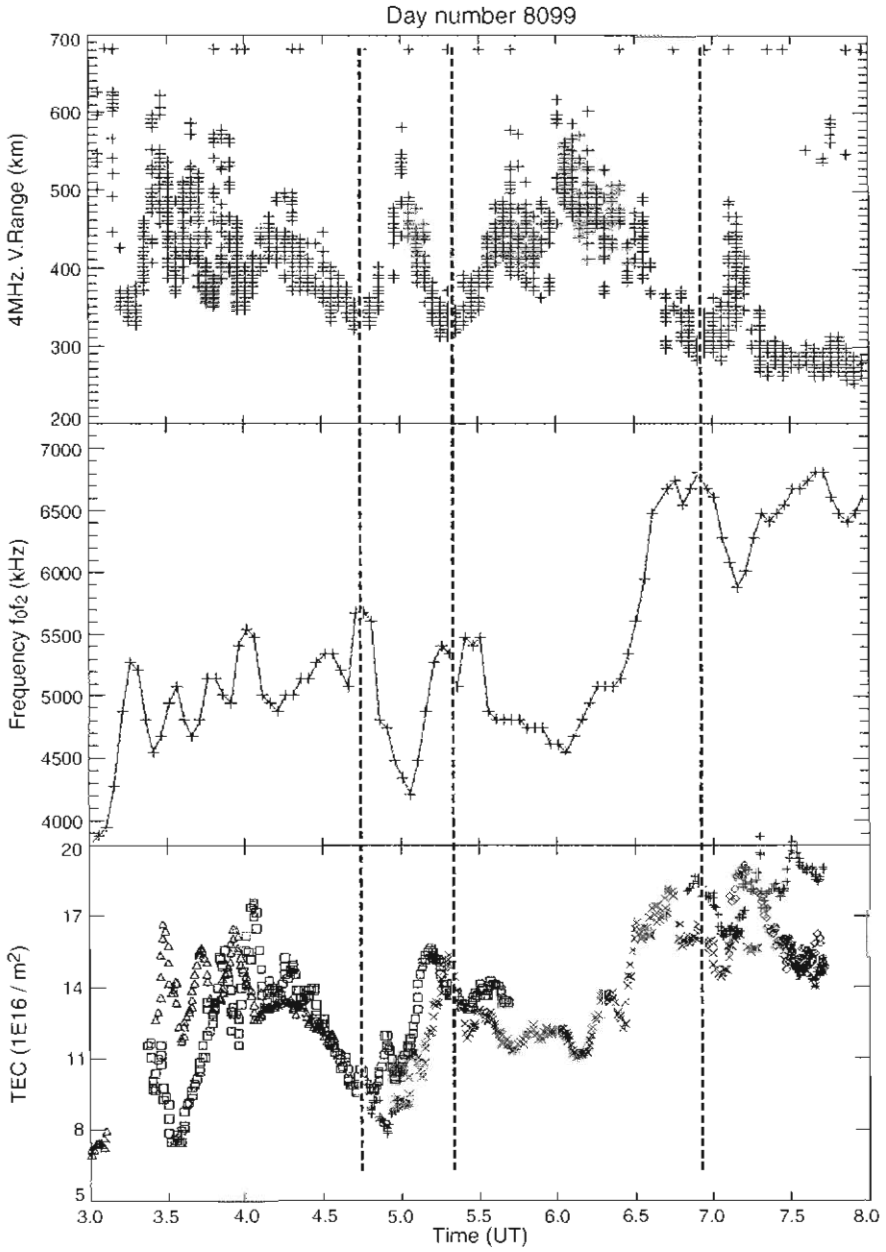


Figure 2. *A comparison between patches identified on the 4 MHz h't plot, f_0F_2 critical frequencies and TEC from GPS satellites (different symbol for each satellite) for 9 April 1998.*

north of Casey). Some time discrepancies exist in the observations of these increases and the closest approach of the patches, most significantly with the earlier patch (some 20 minutes), but also to a lesser extent with the second patch.

A simple explanation for this would be the physical separation of the observations, with GPS TEC measurements corresponding to some 2–3 degrees north (geographic) of the h't and f_oF_2 measurements. Assuming a separation of around 3 degrees latitude, the 20 minute delay of the first patch would correspond to a patch velocity of around 300 m s^{-1}. Horizontal drift velocities derived from DPS drift measurements decreased considerably (from around 600 m s^{-1} to less than 200 m s^{-1}) for approximately 20 minutes following a sudden change in IMF B_y/B_z at 0400 UT, evident in Figure 3.

A second possible explanation for the absence of a direct correlation between observations involves the sudden changes in the IMF B_y and B_z occurring just prior to appearance of the first patch (see Figure 3). It is possible that these changes may have altered or disrupted the

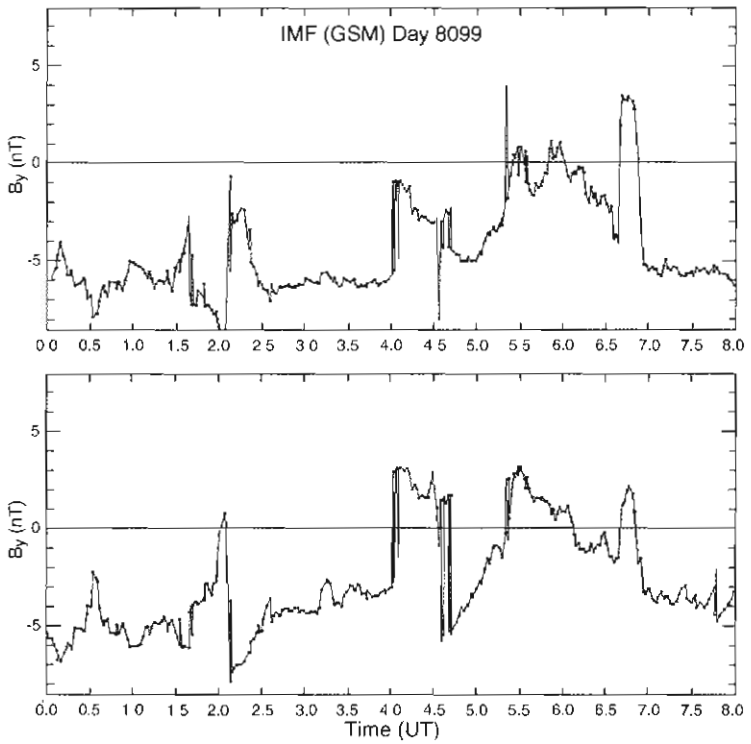


Figure 3. *Interplanetary magnetic field parameters B_y and B_z for 9 April 1998.*

polar circulation pattern sufficiently for the patch to miss the TEC measurement points. DPS measurements around 0400 UT indicate a drift direction shifted approximately 45 degrees away from the antisunward (geographic south) direction towards the post-midnight sector (i.e. originating in the afternoon sector).

An example of the limited spatial extent of some patches is demonstrated later in this paper. It is interesting to note as well that the time interval between the first two identified patches (approximately 40 minutes) appears to correspond to the time between the preceding

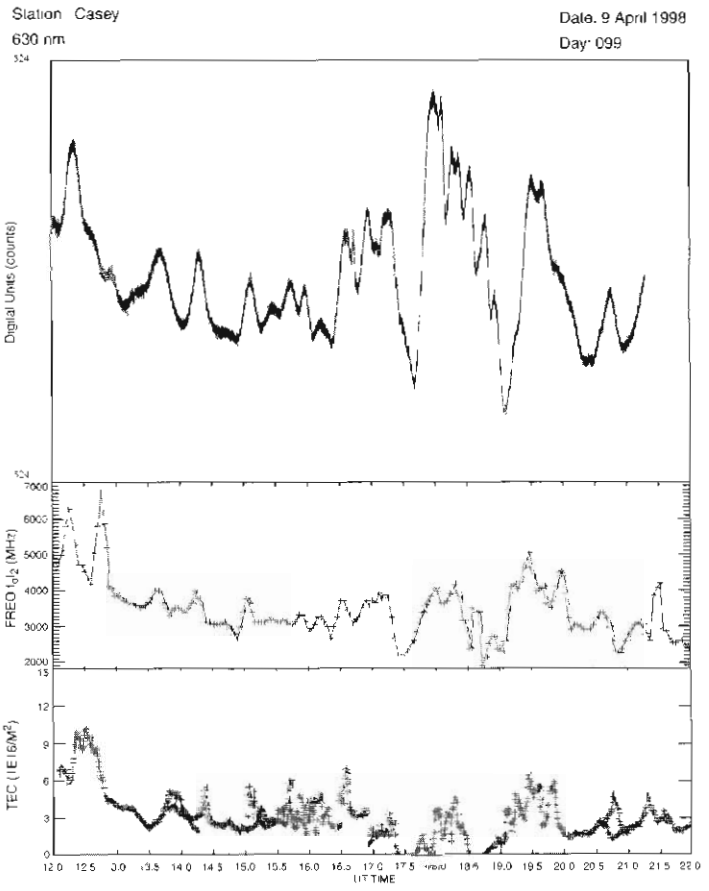


Figure 4. *A series of patches seen in 630 nm photometer data, compared with f_0F_2 critical frequencies and TEC measurements for 9 April 1998.*

south-north and north-south transitions in B_Z suggesting a possible connection.

Figures 4 and 5 show a period later on the same day when optical measurements were available from a 630 nm wide-angle photometer. This photometer has a 60 degree field-of-view and is colocated with the digisonde. Hence peaks in 630 nm airglow arising from patches should correspond temporally to peaks in f_0F_2 critical frequency. Figure 5 shows the state of the IMF during the period, indicating a generally southward (negative) B_Z component conducive to polar patch formation (Buchau *et al.*, 1983). Figure 4 includes the photometer measurements (uncalibrated), f_0F_2 critical frequency and TEC. This figure shows almost a continuous series of enhancements (patches) apparent in the 630 nm

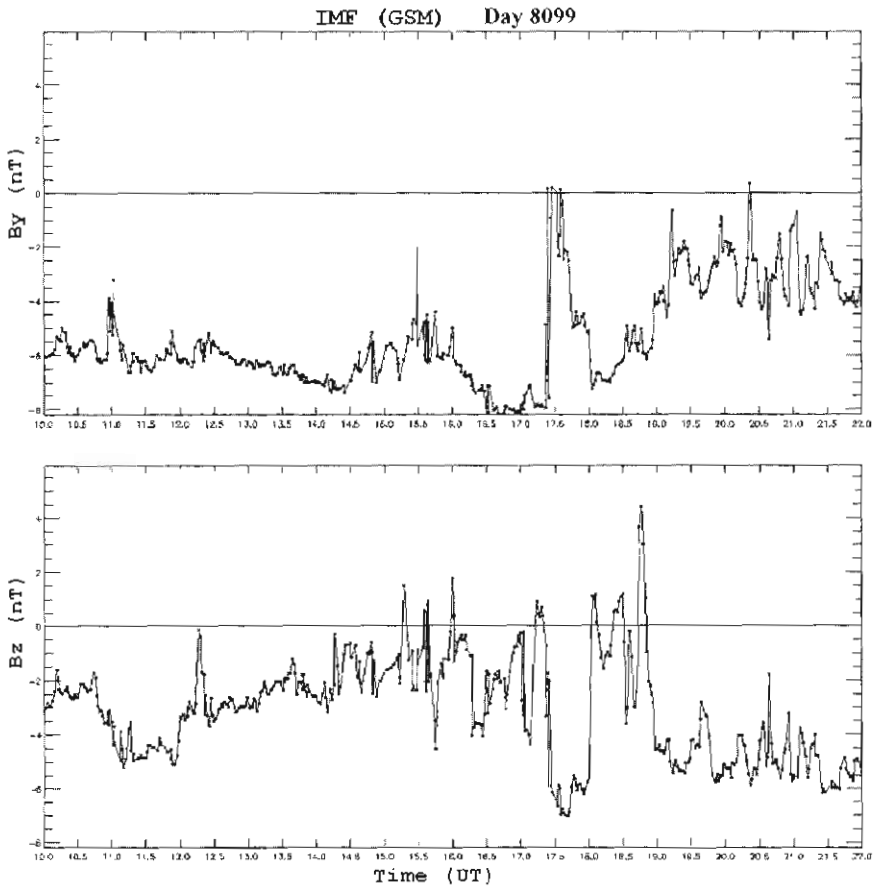


Figure 5. *Interplanetary magnetic field parameters B_y and B_Z for 9 April 1998.*

band. These generally correspond well with both f_oF_2 frequencies and TEC measurements, although there are large differences in the relative magnitudes of the increases as seen by the different techniques. This is in line with the results of Sojka *et al.* (1997).

13.3.2 10 April 1998

Figure 6 shows a h't plot at 4 MHz on 10 April 1998 (day 100). Again the characteristic U-shaped traces are seen indicating the presence of large patch features. In this plot, colour is used to show basic Doppler shift in the top panel and direction of arrival in the bottom panel. The change in Doppler can be clearly seen in the top panel as the patch passes the point of closest approach, from the positive Doppler (cyan) of the approaching patch to the negative Doppler (red) of the receding patch. A signature is not so easily seen in the bottom panel as most of the echoes have been classified as overhead (Vo) within 20 degrees of zenith.

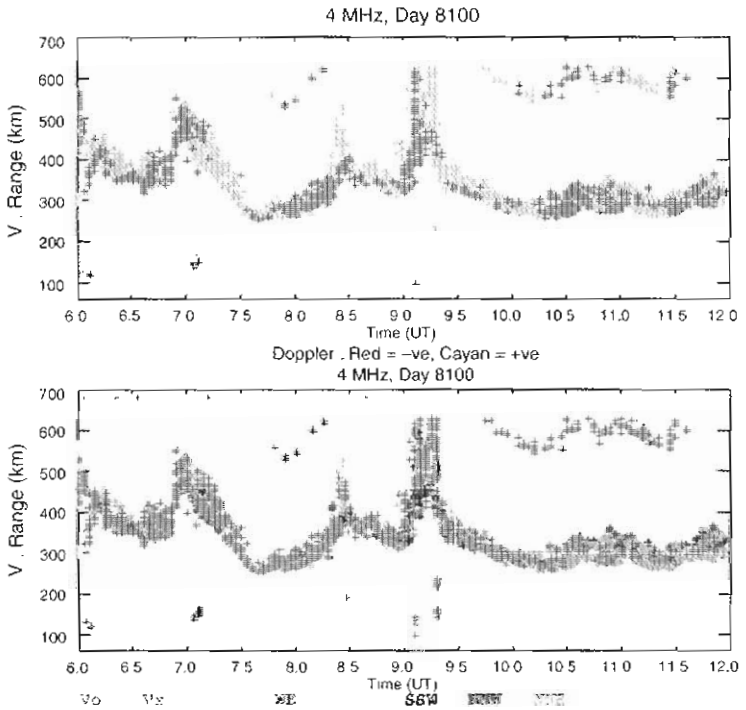


Figure 6. *Virtual range versus time plots at 4 MHz frequency, showing Doppler shifts and directions of arrival for a series of patches identified on 10 April 1998.*

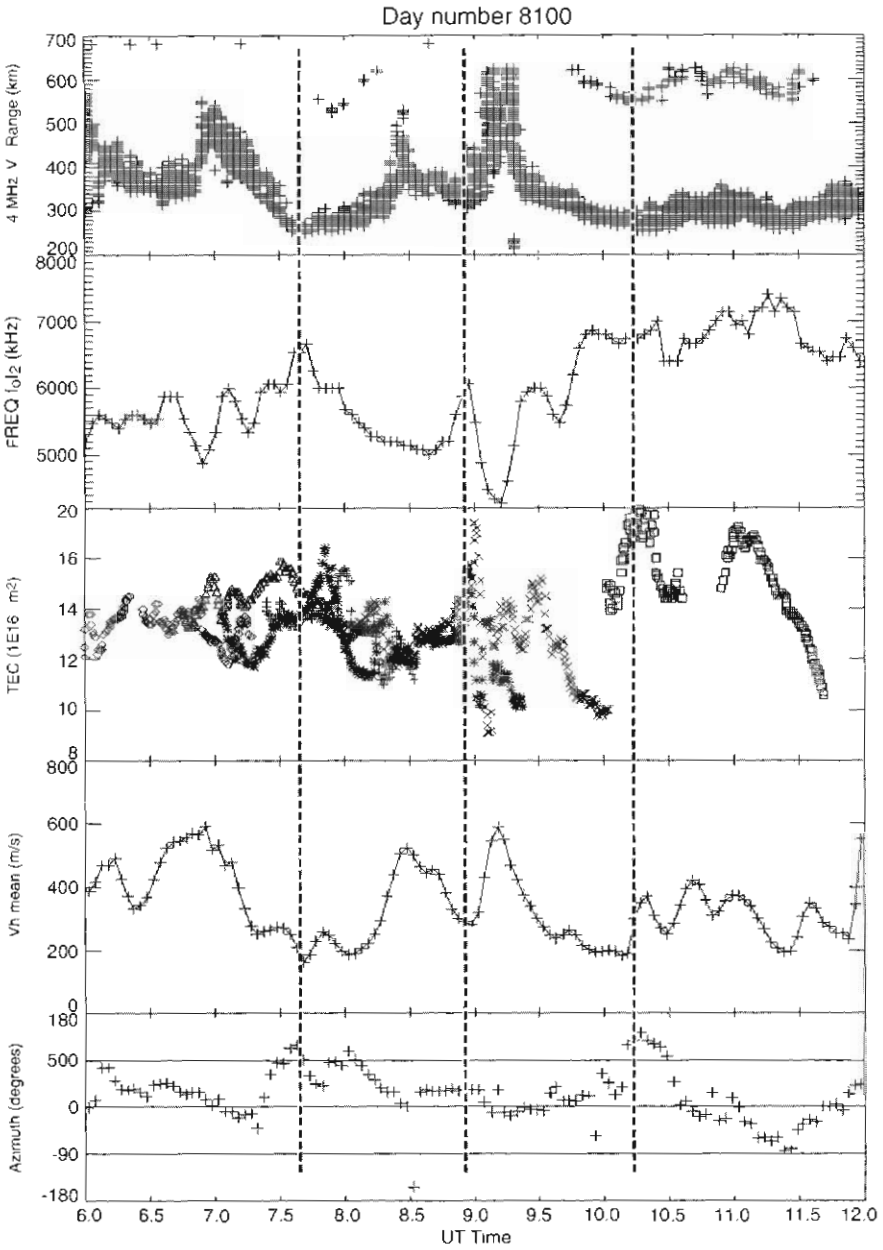


Figure 7. *A comparison between patches identified on the 4 MHz h't plot, f_oF_2 critical frequencies, TEC measurements (different symbols for each satellite), horizontal drift velocity and plasma drift direction (referenced to the antisunward direction) for 10 April 1998.*

Figure 7 shows the same patch sequence in h't, with corresponding critical frequencies and TEC measurements. Also included on this plot is the mean horizontal drift velocity determined from DPS drift measurements and the drift azimuth (the apparent direction of plasma motion). The centre line (0 degrees) in this lower panel represents the antisunward direction. The vertical dashed lines again indicate the approximate times of closest approach of the patches. The identification of an unambiguous single patch becomes difficult in the period after 0930 UT, with f_oF_2 and TEC indicating the possible presence of two or more overlapping patches. What can be seen from this figure is again the close relationship between the U-shaped feature and peaks in f_oF_2 and TEC. The fourth panel of this figure also indicates an apparent link between the patches and horizontal drift velocity, with minima observed at the points of closest approach to the station and surges in velocity associated with the patch edges. This is in agreement with some of the findings of MacDougall *et al.* (1996). From this figure there appears to be no conclusive signatures of patches in the drift azimuth directions. The first patch in the series is associated with a 90 degrees (dawnward) fluctuation in drift azimuth, while the second patch appears to have generally antisunward flow. The third patch and associated drift directions could possibly be described as showing a vortical motion (i.e. drift azimuth swinging from +90 degrees through antisunward to -90 degrees).

13.3.3 16 April 1998

The last case study is shown in Figure 8, again in the form of a h't plot. As with the previous examples, this is compared with f_oF_2 frequencies, TEC and horizontal drift velocity (smoothed in this plot with a three point running mean filter). Up to six patches are indicated by the vertical dashed lines. This period differs significantly from the previous cases in that up until approximately 0730 UT the IMF B_Z component is consistently northward (positive) (see Figure 9). As a consequence of this, the drift azimuths tend to show sunward flow, although in general the directions are quite variable, up to this time. At 0730 UT as B_Z turns southward (negative), the drift directions appear to stabilise and swing more antisunward, although at times showing up to 90 degrees deviation as B_Z oscillates between northward and southward. As shown on Figure 8, there is again a close correlation between the U-shaped patch signatures, peaks in f_oF_2 and TEC (although there was a limited amount of this data available on this day), and minima in horizontal drift velocity. The TEC data available (shown in more detail in Figure 10) displays the apparent narrow longitudinal extent of the final patch feature. This figure shows the TEC as measured along the ray paths to the two GPS satellites visible at the time. A large increase in TEC (from 10–20 TEC units) is seen on satellite PRN26 at a sub-ionospheric longitude of 109.2°E, while no similar increase is seen on satellite PRN2 at a sub-ionospheric longitude of 112.5°E. At this time of day (1000–1100 UT, late afternoon local time), patches would be expected to approach from the west (geographic) (if travelling approximately antisunward) and hence would tend to be observed on a more westward satellite as found in this

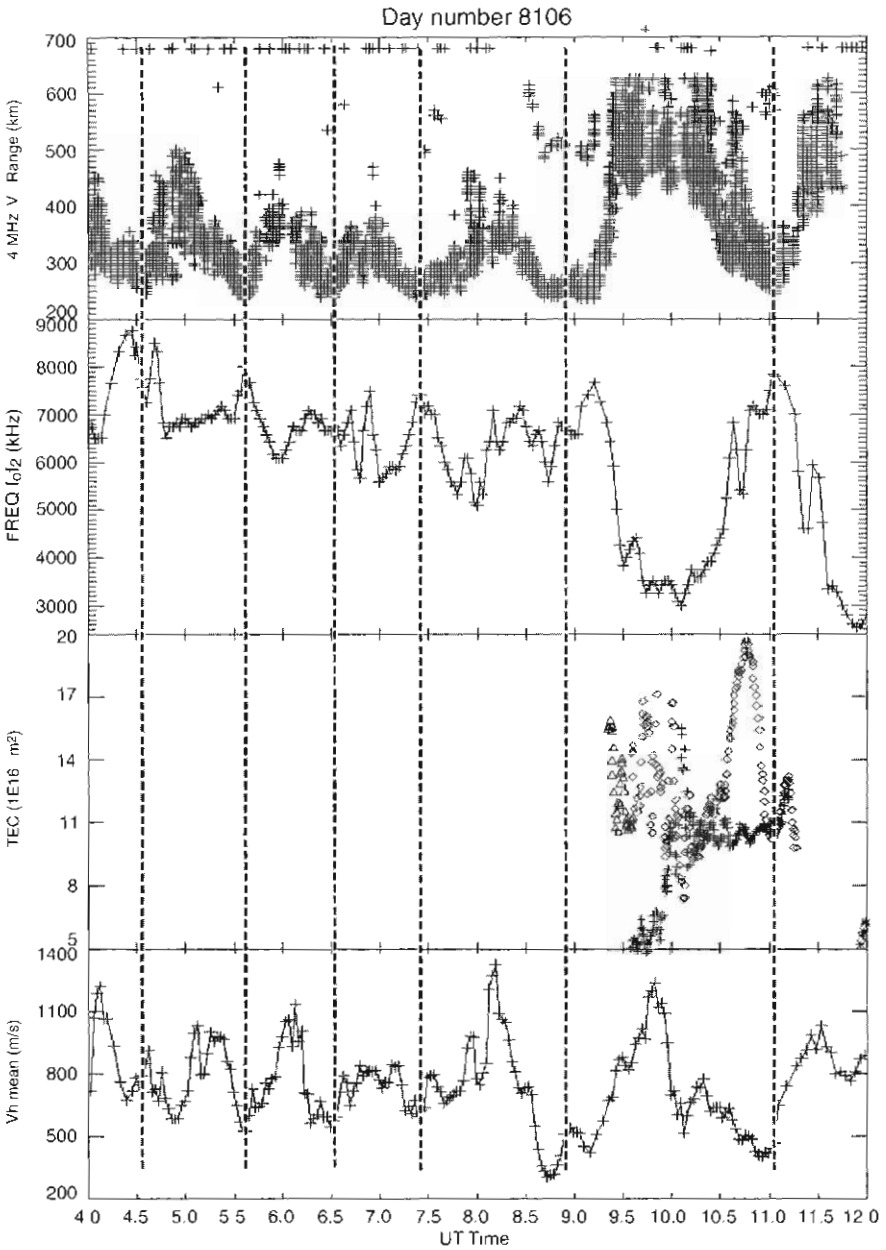


Figure 8. *A comparison between patches identified on the 4 MHz h't plot, f_0F_2 critical frequencies, TEC measurements (different symbols for each satellite) and horizontal drift velocity for 16 April 1998.*

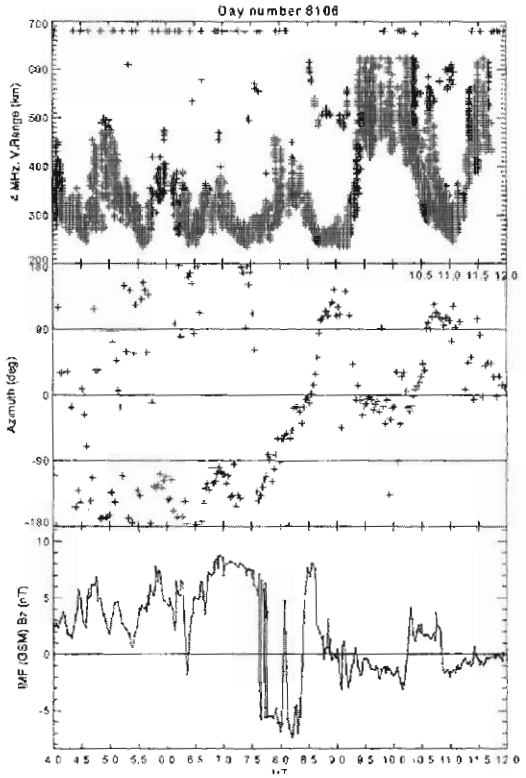


Figure 9. *A comparison between patches identified on the 4 MHz h't plot, plasma drift direction (referenced to the antisunward direction) and IMF B_Z component for 16 April 1998.*

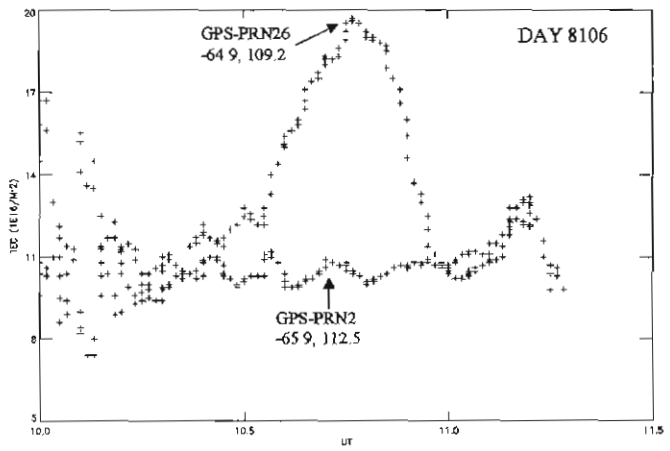


Figure 10. *TEC measurements from GPS satellites for a single patch identified in Figures 10 and 11, showing positions and measurements from the two individual satellites visible at the time for 16 April 1998.*

case. However, the complete lack of any increase on the second satellite, only approximately 150 km away, suggests a sudden, large density gradient at the edge of this particular patch. This is also suggested by the large variation seen in f_0F_2 in Figure 8.

13.4 Discussion and conclusions

Although only preliminary at this stage, the results from these case studies successfully demonstrate the use of the DPS-4 digisonde in identifying and studying polar cap ionosphere features such as polar patches. The polar patches have been identified using a U-shaped signature found in northern-hemisphere studies by James and MacDougall (1997). The patches have been compared and correlated with ionogram critical frequencies, Doppler shifts and directions, Total Electron Content from GPS satellite observations, plasma drift velocity and direction (also from DPS-4 measurements), 630 nm photometer measurements and Interplanetary Magnetic Field parameters.

The significant finding from these case studies is the apparent correlation of horizontal drift velocities with patches. Surges in the observed velocity are associated with the patch edges. No definite conclusions are made regarding this phenomenon at this time due to the possibility that these velocity surges are not part of the bulk plasma flow, but may in fact be an 'artifact' caused by turbulence or instabilities at the patch edges (Rodger, 1998). However, as one of the main creation mechanisms for polar patches are Flow Channel Events (fast plasma jets), it seems reasonable to expect remnants of such events to be observable in the patches (in particular, as most of the patches shown in this study are on the dayside, close to the cusp region where they are formed).

With further study and analysis (incorporating scintillation measurements, SuperDARN radar data and other ionosondes) we expect to be able to resolve some of the questions left unanswered in the present study.

13.5 Acknowlegements

R.P. Lepping and the WIND MAG (MFI) Data Processing team of NASA/Goddard Space Flight Center are thanked for allowing us to use WIND IMF data. The Australian Surveying and Land Information Group (AUSLIG) and La Trobe University are thanked for the use of data from their GPS receivers located at Casey.

References

Anderson, D.N., Buchau, J. and Heelis, R.A. (1998). Origin of density enhancements in the winter polar cap ionosphere. *Radio Science*, 23: 513–519.

Breed, A.M., Goodwin, G.L. and Silby, J.H. (1998). Total electron content measurements in the southern hemisphere using GPS satellites, 1991 to 1995. *Radio Science*, 33: 1705–1726.

Breed, A.M., Maddern, T.M., Dyson, P.L. and Morris, R.J. (2001). Dynamics of the polar cap ionosphere Part 1: applications of digisonde animations in ionospheric research. In: Morris, R.J. and Wilkinson, P.J. (Eds.). *ANARE Reports 146*. Australian Antarctic Division. Pp. 181–188.

Buchau, J., Reinisch, B.W., Weber, E.J. and Moore, J.G. (1983). Structure and dynamics of the winter polar cap F-region. *Radio Science*, 18: 995–1010.

Buchau, J., Weber, E.J., Anderson, D.N., Carlson, H.C., Moore, J.G., Reinisch, B.W. and Livingston, R.C. (1985). Ionospheric structures in the polar cap: their origin and relation to 250 MHz scintillation. *Radio Science*, 20: 325–338.

Coley, W.R. and Heelis, R.A. (1995). Adaptive identification and characterisation of polar ionization patches. *Journal of Geophysical Research*, 100: 23819–23827.

Crowley, G. (1996). Ionospheric patches and blobs. In: Ross, W. (Ed.). *The Review of Radio Science 1993–1996*. Oxford University Press. Pp. 619–649.

Fukui, K., Buchau, J. and Valladares, C.E. (1994). Convection of polar cap patches observed at Qaanaaq, Greenland during the winter of 1989–1990. *Radio Science*, 29: 231–248.

Hill, J.R. (1963). Sudden enhancements of F-layer ionization in the polar regions. *Journal of Atmospheric Science*, 20: 492–497.

James, H.G. and MacDougall, J.W. (1997). Signatures of polar cap patches in ground ionosonde data. *Radio Science*, 32: 497–513.

Kivanc, Ö. and Heelis, R.A. (1997). Structures in ionospheric number density and velocity associated with polar cap ionization patches. *Journal of Geophysical Research*, 102: 307–318.

Lockwood, M. and Carlson, H.C. (1992). Production of polar cap electron density patches by transient magnetopause reconnection. *Geophysical Research Letters*, 19: 1731–1734.

MacDougall, J.W., Grant, I.F. and Hamza, A. (1996). Velocity fluctuations associated with polar patches. *Radio Science*, 31: 595–606.

McEwen, D.J. and Harris, D.P. (1995a). Observations of F-layer patches and their convection over the north magnetic pole. *Advances in Space Research*, 16: 69–72.

McEwen, D.J., Harris, D.P., MacDougall, J.W. and Grant, I.F. (1995b). Drifting F-layer patches over the magnetic pole. *Journal of Geomagnetism and Geoelectricity*, 47: 527–537.

Nichols, J., Hansen, A., Walter, T. and Enge, P. (1999). High latitude measurements of ionospheric scintillation using the NSTB. Department of Aeronautics and Astronautics, Stanford University (draft paper).

Ogawa, T., Nishitani, N., Pinnock, M., Sato, N., Yamagishi, H. and Yukimatu, A.S. (1998). Polar cap patches and auroral blobs observed with Antarctic HF radars: preliminary results. *Proceedings NIPR Symposium on Upper Atmospheric Physics*, 11: 11–18.

Ogawa, T., Nishitani, N., Pinnock, M., Sato, N., Yamagishi, H. and Yukimatu, A.S. (1998). Antarctic HF radar observations of irregularities associated with polar patches and auroral blobs: a case study. *Journal of Geophysical Research*, 103: 26547–26558.

Parkinson, M.L., Breed, A.M., Dyson, P.L. and Morris, R.J. (1999a). Application of the Dopplionogram to Doppler-sorted interferometry measurements of ionospheric velocity. *Radio Science*, 34: 899–912.

Parkinson, M.L., Breed, A.M., Dyson, P.L. and Morris, R.J. (1999b). Signatures of the ionospheric cusp in digital ionosonde measurements of plasma drift above Casey, Antarctica. *Journal of Geophysical Research*, 104: 22487–22498.

Reinisch, B.W., Bullett, T.W., Scali, J.L. and Haines, D.M. (1995). High latitude digisonde measurements and their relevance to IRI. *Advances in Space Research*, 16: 17–26.

Rodger, A.S., Pinnock, M., Dudeney, J.R., Baker, K.B. and Greenwald, R.A. (1994a). A new mechanism for polar patch formation. *Journal of Geophysical Research*, 99: 6425–6436.

Rodger, A.S., Pinnock, M., Dudeney, J.R., Waterman, J., de la Beaujardiere, O. and Baker, K.B. (1994b). Simultaneous two hemisphere observations of the presence of polar patches in the nightside ionosphere. *Annales Geophysicae*, 12: 642–648.

Rodger, A.S. and Graham, A.C. (1996). Diurnal and seasonal occurrence of polar patches. *Annales Geophysicae*, 14: 533–537.

Rodger, A.S. (1998). Polar patches–Outstanding issues. In: Moen *et al.* (Eds.). *Polar Cap Boundary Phenomena*. Kluwer Academic Publishers. Pp. 281–288.

Rosenberg, T.J., Wang, Z., Rodger, A.S., Dudeney, J.R. and Baker, K.B. (1993). Imaging riometer and HF radar measurements of drifting F-region electron density structures in the polar cap. *Journal of Geophysical Research*, 98: 7757–7764.

Scali, J., Reinisch, B., Dozois, C., Bibl, K., Kitrosser, D., Haines, M. and Bullett, T. (1995). *Digisonde drift analysis*. University of Massachusetts, Lowell, Center for Atmospheric Research.

Sojka, J.J., Bowline, M.D., Schunk, R.W., Decker, D.T., Valladares, C.E., Sheehan, R., Anderson, D.N. and Heelis, R.A. (1993). Modelling polar cap F-Region patches using time varying convection. *Geophysical Research Letters*, 20: 1783–1786.

Sojka, J.J., Schunk, R.W., Bowline, M.D. and Crain, D.J. (1997). Ambiguity in identification of polar cap F-region patches: contrasting radio and optical observation techniques. *Journal of Atmospheric and Solar Terrestrial Physics*, 59: 249–258.

Steele, D.P. and Cogger, L.L. (1996). Polar patches and the 'tongue of ionization'. *Radio Science*, 31: 667–677.

Valladares, C.E., Basu, S., Buchau, J. and Friis-Christensen, E. (1994). Experimental evidence for the formation and entry of patches into the polar cap. *Radio Science*, 29: 167–194.

Valladares, C.E., Decker, D.T., Sheehan, R. and Anderson, D.N. (1996). Modelling the formation of polar cap patches using large plasma flows. *Radio Science*, 31: 573–593.

Valladares, C.E., Decker, D.T., Sheehan, R., Anderson, D.N., Bullett, T. and Reinisch, B.W. (1998). Formation of polar cap patches associated with north-to-south transitions of the interplanetary magnetic field. *Journal of Geophysical Research*, 103: 14657–14670.

Wang, Z., Rosenberg, T.J., Stauning, P., Basu, S. and Crowley, G. (1994). Calculations of riometer absorption associated with F-region plasma structures based on Sondre Stromfjord incoherent scatter radar observations. *Radio Science*, 29: 209–215.

Weber, E.J., Buchau, J., Moore, J.G., Sharber, J.R., Livingston, R.C., Winningham, J.D. and Reinisch, B.W. (1984). F layer ionization patches in the polar cap. *Journal of Geophysical Research*, 89: 1683–1694.

Weber, E.J., Klobuchar, J.A., Buchau, J., Carlson, H.C., Livingston, R.C., de la Beaujardiere, O., McCready, M., Moore, J.G. and Bishop, G.J. (1986). Polar cap F-layer patches: structure and dynamics. *Journal of Geophysical Research*, 91: 12121–12129.

14. ON THE DETERMINATION OF VERTICAL PROFILES OF IONOSPHERIC VELOCITY FROM DIGITAL IONOSONDE MEASUREMENTS

P.L. Dyson[1], M.L. Parkinson[1], D.P. Monselesan[2]* and R.J. Morris[2]

[1] Department of Physics
La Trobe University
Victoria 3083 Australia
(email: peter.dyson@latrobe.edu.au)

[2] Atmospheric and Space Physics
Australian Antarctic Division
Kingston Tasmania 7050 Australia

* Now at IPS Radio and Space Services

Abstract

The so-called ionosonde drift technique of determining ionospheric velocities from angle-of-arrival and Doppler shift measurements is now well established. In the standard digital ionosonde technique, commonly called the drift mode, the velocity of the F-region is usually determined without detail of its height variation. In this paper, we report the measurement of the height variation of the velocity profile through the E- and F-regions at Casey, Antarctica. In the standard analysis, velocities are interpreted as if they arose from mirror reflectors in free space. According to Dyson (1975), this gives valid results when the ionosphere moves with a constant bulk velocity. However, the velocity may vary through the ionosphere and the purpose of this paper is to discuss the affect of such variations on the derivation of the drift velocity. A simplified inversion routine is developed enabling average drift velocity profiles to be corrected for the velocity variation with height. The procedure has been applied to an average daytime velocity height profile obtained at Casey during a four day campaign in March 1996. The drift velocity component perpendicular to the magnetic field direction was found to increase rapidly through the bottom of the E-region from a value of 170 m s^{-1} at 95 km to a peak value of over 800 m s^{-1} at 110 km. It then dropped rapidly to around 700 m s^{-1} at 115 km and remained in the range 600–700 m s^{-1} up into the F-region. The corrected values were about 10% larger in the lower E-region but the difference is not important because of the very rapid change in velocity with height. In the F-region the correction was 3% or less. At 110 km where a narrow, intense velocity peak occurred, the correction was 20%, dropping to 7.5% near 150 km. Even in the region of largest correction, the change in the shape of the velocity profile was modest. The results support the interpretation of drift velocities as reflection point velocities (at least within ~ 5–10%) except in regions of extreme velocity

change and certainly support the routine interpretation of high-latitude measurements of the F-region velocity as a bulk convection velocity.

14.1 Introduction

The advent of digital ionosondes has brought two new major developments to ionospheric sounding. The first is automatic scaling of echo traces to derive ionospheric parameters and true height profiles of electron density. The second is the combination of angle-of-arrival and Doppler shift measurements to routinely determine ionospheric motions.

The determination of true height profiles involves inversion of the equation

$$h' = \int_0^{h_R} \mu' dh \tag{1}$$

where h' is the virtual height,

μ' is the group refractive index, and

h_R is the height of reflection.

Sophisticated inversion methods have been developed to account for various difficulties that arise in, for example, determining a starting height and accounting for a valley between the E- and F-regions (e.g. Titheridge, 1985; Gamache *et al.*, 1992).

The Doppler shift, Δf, of a radio signal transmitted at frequency, f, is given by Bennett (1968) as

$$\Delta f = -\frac{f}{c}\frac{dP}{dt} = -\frac{f}{c}\left(\int_a^b \frac{\partial \mu}{\partial t}\cos\alpha\, ds + \left[\mu \hat{\boldsymbol{p}}.\boldsymbol{v}\right]_a^b\right) \tag{2}$$

where

μ is the phase refractive index,

α is the angle between the wave normal and ray directions,

v is the velocity at the ray end points, a and b,

$\hat{\boldsymbol{p}}$ is a unit vector in the wave normal direction,

ds is an element of arc length, and

c is the speed of light in free space.

If the Doppler shift is due to a reflector in free space ($\mu = 1$) returning a signal to its origin, then Equation (2) reduces to

$$\Delta f = -\frac{f}{c}\frac{dP}{dt} = -\frac{2f}{c}\,\boldsymbol{v}_R.\hat{\boldsymbol{p}} = -\frac{2f}{c}\left(\hat{\boldsymbol{v}}_R.\hat{\boldsymbol{p}}\right)v_R \tag{3}$$

where $\boldsymbol{v}_R$ is the velocity of the reflector and $\hat{v}_R$ is the corresponding unit vector, $\hat{\boldsymbol{p}}$ is the unit vector in the ray direction. The factor of two comes about because the ray path is from the transmitter to the reflector and back again. This result incorporates Pfister's result for a reflector in free space moving horizontally (Pfister, 1971).

Now $\boldsymbol{v}_R.\hat{\boldsymbol{p}}$ implies the component of velocity along the angle-of-arrival direction. Hence this component of the reflector's velocity can be

determined from angle-of-arrival and Doppler shift measurements. If at least three echoes at different angles-of-arrival are obtained from reflectors moving with the same velocity, then the full velocity vector of the reflectors can be determined.

Dyson (1975) showed that Equation (3) is not restricted to just the case of a reflector in free space but applies to ionospheric echoes returned to a transmitter site when the ionosphere moves with a single bulk velocity. This point was further developed by Bennett and Dyson (1993) who showed that the result may be regarded as a generalisation of Pfister's result for a moving reflecting surface in free space (Pfister, 1971). In practice then, digital ionosonde measurements are used to determine the ionospheric motion by applying Equation (3) to concurrent measurements of Doppler shift and angle-of-arrival of many near-vertical oblique echoes (e.g. Scali *et al.*, 1995). The technique is now commonly referred to as the 'drift mode' and the velocity so obtained, the 'drift velocity'. This convention is adopted here and we specifically use the term 'drift velocity' for velocities derived using Equation (3). That is, if the ionosphere moves with a single bulk velocity then the measured drift velocity, $v_D = v_R$. Of course it should be remembered that the motion of the ionosphere is affected by a number of processes, such as convection driven by electric fields, passing TID's, and coupling of the neutral wind. They will all affect the drift velocity though in certain situations one may dominate.

Digital ionosondes have the capability of observing the variation of ionospheric velocity with height by determining the drift velocity as the transmitting frequency is swept in a similar way to that used to obtain a swept-frequency ionogram. However, obtaining a good quality velocity determination using the drift technique requires transmission on a single frequency for at least several seconds. Consequently, constraints on the use of the radio spectrum, and the amount of data processing involved, mean that usually only a single velocity is obtained in routine drift determinations of ionospheric velocity.

In spite of these constraints, it has been possible to run some special campaigns at Casey station, Antarctica, to obtain the drift velocity as a function of height through the bottomside ionosphere (Parkinson *et al.*, 1997). A feature of the results is that the velocity varies rapidly with height through the E-region. Thus the assumption of a single, uniform ionospheric velocity, a requirement for strict validity of Equation (3), is violated. The significance for the interpretation of the frequency profiles is examined here.

14.2 Observations

The results presented were obtained with a Lowell Digital Portable Sounder-4 (DPS-4) operated at the Australian Antarctic station, Casey (66.3°S, 110.5°E, -80.8° CGM latitude). Whilst operation of other digital sounders may differ in some respects from the DPS-4, the general principles of relevance here are the same.

The DPS-4 (Haines and Reinisch, 1993) uses an array of four receiving antennas and when operated in the drift mode, a time series of the complex amplitude is obtained at each antenna and for each frequency of operation. Each time series is converted to a Doppler frequency spectrum enabling signals with the same Doppler shift to be identified at each antenna. The relative phases of these signals then gives the angle-of -arrival. Hence Equation (3) can be applied to each Doppler signal to determine the ionospheric drift velocity. In practice the ionosphere is not always a smooth reflector and multiple echoes are observed simultaneously but with different Doppler shifts and the ionospheric drift velocity is determined by the least-squares application of Equation (3). Details of this approach have been given by Scali *et al.* (1995) who have implemented the technique as part of the standard analysis software of the DPS-4.

At a high-latitude station such as Casey, the ionosphere is often very irregular and many echoes are observed. This improves the accuracy of the drift velocities. Furthermore, the F-region motion at high-latitudes is dominated by convection, simplifying the interpretation of the drift velocities. Comparisons with other techniques (Scali *et al.*, 1995; Grant *et al.*, 1995; Smith *et al.*, 1998) have shown that the drift-mode technique is reliable at high-latitudes.

During January–March 1996 several special experimental campaigns were conducted using the Casey DPS-4 to compare drift velocities in the E- and F-regions and to obtain vertical profiles of the drift velocity. The detailed results of those campaigns are discussed elsewhere (Parkinson *et al.*, 1997). Of relevance here is the observation that the drift velocity varied with height, particularly through the E-region, and the consequence of this on the interpretation of drift velocities is discussed in detail.

As convection dominates at Casey it is convenient to consider the drift velocity component perpendicular to the magnetic field lines and this component will be referred to here as the convection velocity. Since the dip angle at Casey is 82°, the convection velocity is not significantly different from the horizontal velocity component. Figure 1 shows an average daytime velocity profile for the convection velocity derived by applying Equation (3) to measurements obtained within three hours of magnetic noon during 11–14 March 1996. The corresponding average true height profile is also shown. Further details on the analysis methods used to obtain these profiles were given by Parkinson *et al.* (1997).

Note that in Figure 1 the velocity increases rapidly by several hundred m s^{-1} between 95 and 110 km. The sharp enhancement of velocity near the E-layer peak is an artifact caused by averaging over the non-stationary random data samples. The small number of echoes detected from the E-layer peak occurred during intervals when the convection velocity was unusually large. Within the F-region the velocity remains reasonably steady up to the height of the F_2 peak, with a value in the range 650–700 m s^{-1}. Dyson (1975) showed that the drift velocity derived

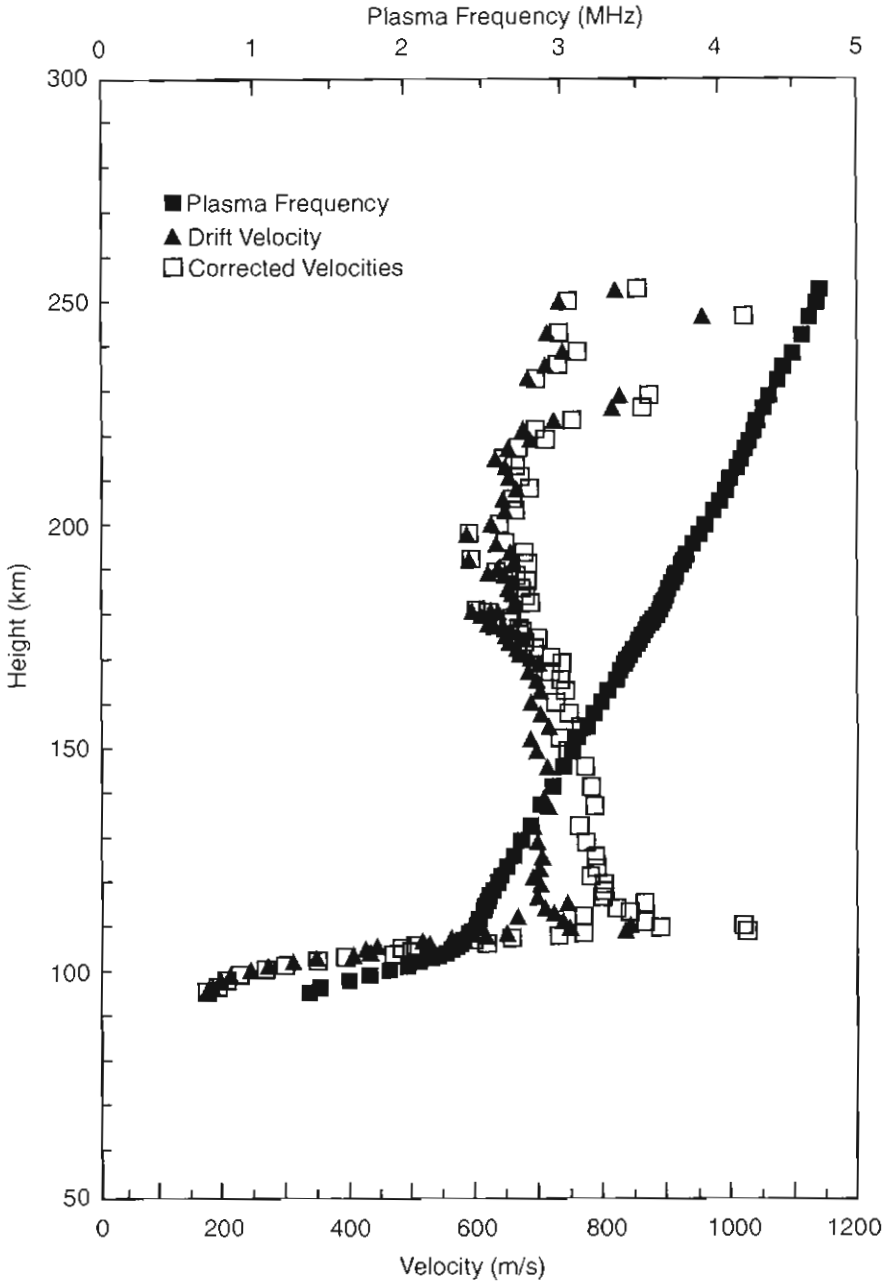


Figure 1. *Shown are: (i) the average height profile of convection velocity measured at Casey (triangles). The profile was obtained using measurements made within three hours of magnetic noon during 11–14 March 1996; (ii) the corresponding average plasma frequency profile is also shown (filled squares); and (iii) the velocities derived for the reflection points of the transmitted frequencies (open squares).*

using Equation (3) will underestimate the velocity at the reflection height if the velocity in the ionosphere increases with height. That is, the drift velocity will be less than the true velocity at the reflection height. In the simple, illustrative example in Dyson (1975), in which the vertical velocity varied linearly with height from a value of zero at the base of the ionosphere, the drift velocity was half the value of the true velocity at the reflection point. Hence we can anticipate that the drift velocities between 95 and 110 km will be underestimates of the velocities at those heights. If it were real, the sharp peak in velocity at 110 km would also affect the determination of velocities above that height. However, well into the F-region there is little variation in velocity with height so the drift velocities should be accurate there. This is examined quantitatively in the next section.

14.3 Analysis

The aim is to determine the effect of the height variation of velocity on the derivation of velocity height profiles using Equation (3). In principle an inversion procedure based on individual measurements could be developed using Equations (1) and (2) to determine both the electron density and velocity profiles. However, since echoes are received over a range of angles-of-arrival, three-dimensional ray tracing would be needed to provide a completely general solution to the problem. Nevertheless it is profitable to proceed in a slightly less general way.

First consider the inversion of Equation (1) to obtain electron density profiles. The methods in common use assume that the ionosphere is horizontally stratified. This is rarely strictly true but the consequential errors in determining the layer structure of the ionosphere are considered to be small, particularly when compared to the likely errors associated with determining the ionospheric base height and E-F valley structure. Inversion in the presence of complicated ionospheric structures has been attempted, for example, Dyson and Benson (1978) determined the electron density variation along ionospheric bubbles by inverting topside ionogram echo traces arising from field-aligned propagation. In that instance they assumed that the electron density contours were perpendicular to the field line within the central part of the bubble itself.

The application of Equation (2) to determine ionospheric velocity vectors depends in fact, on off-vertical reflections occurring. The reason is that the horizontal component of velocity will not contribute to the Doppler shift of a ray propagating at vertical incidence in a horizontally stratified ionosphere, but it will contribute to the Doppler shift of off-vertical rays. Hence the inversion of Equation (2) is inherently more difficult because off-vertical rays must be considered.

14.3.1 Reflection from the $\mu = 0$ contour

To simplify the problem, we assume that for each echo the tilt of the refractive index contours along the ray path is not significantly different to that at the reflection point. If we also ignore the effect of the Earth's magnetic field on the refractive index ($\mu = 0$), then the ray paths will

not deviate significantly from straight lines tilted to the vertical. Since the ionosonde is stationary on the surface of the Earth where $\mu = 1$ and, since under the conditions assumed reflection takes place where $\mu = 0$, Equation (2) becomes

$$\Delta f = -\frac{f}{c}\frac{dP}{dt} = -\frac{2f}{c}\int_S \frac{\partial \mu}{\partial t}\, ds. \tag{4}$$

Suppose the change in μ is due entirely to the motion of the ionosphere, then we may write $\partial\mu/\partial t = v_x(\partial\mu/\partial x) + v_y(\partial\mu/\partial y) + v_z(\partial\mu/\partial z)$ so that, noting that in order to be reflected, μ decreases overall as s increases, Equation (4) becomes

$$\Delta f = -\frac{2f}{c}\left(\hat{v}.\hat{p}\right)\int_0^1 v\, d\mu \tag{5}$$

where $\hat{p}$ is a unit vector in the ray direction.

Hence in this special case the application of the drift technique will yield a drift velocity $v_D = \int_0^1 v d\mu$ rather than $v_D = v_R$ as is the case for motion with a single bulk velocity.

We can now use Equation (5) to determine the velocity at the reflection height provided the ionospheric electron density profile is known (i.e. the spatial variation of μ under the assumptions made). Following the procedure generally used in the inversion of Equation (1) to obtain the electron density profile, we regard the ionosphere as divided into a number of slabs. In this case the slabs are defined by the plasma frequencies at the top and bottom of the slab, and within the slab the true velocity is assumed to vary linearly with the refractive index. Hence the velocity variation in the n^{th} slab is given by

$$v_n = v_{bn} + k_n(\mu - \mu_{bn}) \tag{6}$$

where $k_n = \dfrac{v_{tn} - v_{bn}}{\mu_{tn} - \mu_{bn}}$ is a constant for the n^{th} height interval,

v_{bn} and v_{tn} are the velocities at the bottom and top of the n^{th} slab respectively,

μ_{bn} and μ_{tn} are the refractive indices at the bottom and top of the n^{th} slab respectively.

Hence

$$v_D = \int_0^1 v\, du = \sum_{n=1}^{N}\int_{\mu_{tn}}^{\mu_{bn}} v_n d\mu, \text{ since } \mu_{tn} < \mu_{bn} \tag{7}$$

where N is the number of slabs from the base of the ionosphere to the reflection point.

Evaluating the integral in Equation (7) with the aid of Equation (6) leads to

$$v_D = \frac{1}{2}\sum_{n=1}^{N}\left[\left(v_{bn} + v_{tn}\right)\left(\mu_{bn} - \mu_{tn}\right)\right]. \tag{8}$$

Suppose the drift velocity is measured at a series of o-ray frequencies, f_1, f_2, ... f_N... f_M, to give a series of drift velocities, v_{D1}, v_{D2}, ... v_{DN}, ... v_{DM}. Let the series of frequencies divide the ionosphere into slabs bounded by the same set of plasma frequencies and let the true velocities at the reflection heights of these frequencies be v_{R1}, v_{R2}, ... v_{RN}, ... v_{RM}. The frequency f_1 is reflected at the top of the first slab and from Equation (8) we have $\mu_{b1} = 1$ and $\mu_{t1} = 0$, so

$$v_{R1} = 2v_{D1} - v_B \tag{9}$$

where $v_B = v_{b1}$ is the velocity at the base of the ionosphere.

For the frequency f_2 Equation (8) becomes:

$$v_{D2} = 1/2[(v_B + v_{R1})(1 - \mu_{b2}) + (v_{R1} + v_{R2})(\mu_{b2})], \text{ since } \mu_{t1} = \mu_{b2} \tag{10}$$

giving

$$v_{R2} = \frac{2v_{D2} - (v_{R1} + v_B)(1 - \mu_{b2})}{\mu_{b2}} - v_{R1} . \tag{11}$$

Similarly for the N^{th} frequency,

$$v_{RN} = \frac{2v_{DN} - \sum_{n=1}^{N-1}\left[\left(v_{R(n-1)} + v_{Rn}\right)\left(\mu_{bn} - \mu_{tn}\right)\right]}{\mu_{bN}} - v_{R(N-1)} . \tag{12}$$

Hence the method can proceed step by step using the drift velocity measurements to determine the height profile of the velocity. However, if the slabs represent small steps in μ, the denominator term, μ_{bN}, becomes small and the process becomes sensitive to errors. As a consequence, in application to the velocity profile measurements reported here, this method proved unstable so an alternative procedure was adopted, which used the measured values of v_D as the starting profile for an iterative process.

As a general starting point, suppose $v_E(\mu)$, an estimate of the velocity profile is known, then the true velocity profile may be written as

$$v_R(\mu) = v_E(\mu) + \varepsilon(\mu), \text{ where } \varepsilon(\mu) \text{ is the error.} \tag{13}$$

It follows from Equation (7) that

$$v_D = \int_0^1 (v_E + \varepsilon)\, d\mu = v_{DE} + \int_0^1 \varepsilon\, d\mu \tag{14}$$

where v_{DE} is the drift velocity calculated using Equation (7) applied to $v_E(\mu)$.

Now at each radio frequency we want to determine the velocity at the reflection point and hence it is convenient to write $\Delta\varepsilon(\mu) = \varepsilon(\mu) - \varepsilon_R$ so that

$$v_D = v_{DE} + \varepsilon_R + \int_0^1 \Delta\varepsilon\, d\mu \tag{15}$$

Then to a first approximation, $\varepsilon_R = v_D - v_{DE}$ with the error given by the integral term in the above equation. Now there is no *a priori* guarantee that this error term will be small, but it will obviously be small in regions where the velocity does not change rapidly near the reflection height. Figure 1 shows that this occurs over most of the F-region. The error term will also be small even if the velocity changes rapidly provided the change in μ over the corresponding height range is small. Hence the error term will be small for frequencies reflected in the F-region, despite the rapid change in velocity in the lower E-region.

So, v_D is the set of drift measurements, v_{DE} can be calculated, hence ε_R can be estimated for each v_D using Equation (15), and Equation (13) can then be used to determine each v_R. If necessary the process can be iterated using the newly estimated velocity profile. Since $v_D = v_R$ in the absence of velocity shears, we can take the measured v_D profile as the initial profile.

This procedure has been applied to the average convection velocity profile shown in Figure 1. The results shown are for a single application of the procedure as a second iteration changed the velocities by only a few percent or less. The corresponding average true-height profile, also shown in Figure 1, has been used to specify radio frequencies and to convert the $v(h)$ profile in Figure 1 to $v(\mu)$ profiles. The velocity at the bottom of the ionosphere was taken to be that measured at the lowest height. The corrected velocity values are shown by open squares in Figure 1. It is apparent that while the corrected velocities in the lower E-region are higher (typically 10%) than those measured, the velocity is changing so rapidly with height that the change in the velocity profile is small. The major difference occurs near 110 km where there is a pronounced, narrow peak in the velocity profile. As pointed out earlier, this peak is an artifact of the data. Nevertheless, it is still interesting to examine the effect such a feature has on the determination of velocities. We note that the value of the peak velocity is 20% greater than the corresponding drift velocity and the presence of this peak affects the measurements at heights immediately above, and up to approximately 150 km. Once this altitude is reached, the correction to the drift velocity is less than 7.5%, and is no more than a few percent in the F-region.

Overall then, the analysis shows that the derivation of ionospheric drift velocities using Equation (3) gives reasonably accurate values in the F-region. Even in the lower E-region where the velocity changes rapidly with height, drift velocities can be considered accurate since even if they are in error by 10–15% at the actual reflection height, they give the velocity at a height within a kilometre or so. The data examined showed a narrow velocity peak near 110 km altitude which caused the drift velocities to be as much as 20% below the velocity at the corresponding reflection points. This velocity peak had a significant influence on the drift velocities up to about 150 km altitude causing the magnitude of the velocities at the reflection points to be underestimated. However, the velocity profile shape was not affected significantly.

14.3.2 Reflection from below the $\mu = 0$ level

The above analysis assumed that echoes arose from reflections at the $\mu = 0$ level. Of course echoes may arise below this level so we now consider this situation. Suppose an echo occurs from an irregularity where the background refractive index is μ_I. Then using Equation (2), Equation (5) is replaced by

$$\Delta f = -\frac{f}{c}\frac{dP}{dt} = -\frac{2f}{c}\left(\hat{\mathbf{v}}.\hat{\mathbf{p}}\right)\left(\int_{\mu_I}^{1} v\,d\mu + \mu_I v_I\right). \tag{16}$$

Hence we may write

$$v_D = \int_0^1 v_T\,d\mu - \int_0^{\mu_I} v_T\,d\mu + \mu_I v_I \tag{17}$$

where $v_T(\mu)$ is the true velocity profile and v_I is the true velocity at the echoing point.

Now by the mean value theorem we may write Equation (17) as

$$\begin{aligned} v_D &= \int_0^1 v_T\,d\mu - \bar{v}_T\int_0^{\mu_I} d\mu + \mu_I v_I \\ &= \int_0^1 v_T\,d\mu - \left(\bar{v}_T - v_I\right)\mu_I \end{aligned} \tag{18}$$

where $\bar{v}_T$ has a value between v_I and v_R, the value of v_T at the $\mu = 0$ level.

For convenience we represent the velocity profile determined assuming reflection at the $\mu = 0$ level by $v_R(\mu)$. Then since $v_D(\mu)$ represents the measurements,

$$v_D = \int_0^1 v_T\,d\mu - \left(\bar{v}_T - v_I\right)\mu_I = \int_0^1 v_R\,d\mu. \tag{19}$$

Now if echoes arise from below the $\mu = 0$ level but in a region where the velocity is not changing between the $\mu = \mu_I$ and $\mu = 0$ levels then $\bar{v}_T = v_I$. Hence $\int_0^1 v_T d\mu = \int_0^1 v_R d\mu$ and the application of Equation (5) gives the correct answer. However, if the velocity increases (decreases) between the $\mu = \mu_I$ and $\mu = 0$ levels then $\bar{v}_T > v_I$ ($\bar{v}_T < v_I$), so that $\int_0^1 v_T d\mu > \int_0^1 v_R d\mu$ ($\int_0^1 v_T d\mu < \int_0^1 v_R d\mu$). These inequalities will be true if $v_T > v_R$ ($v_T < v_R$).

The application of the method developed using Equation (5) assumes reflection at the $\mu = 0$ level, so if echoes arise from below this level, errors will occur if the velocity is changing with height. Assuming that the vertical profile of electron concentration is correct, then the velocity v_R will be ascribed to a particular height instead of v_T. Hence if the velocity increases (decreases) with height, the velocity at a particular height will be underestimated (overestimated).

With reference to Figure 1, the major region in which the velocity increases with height (and therefore decreases with $\mu = 0$) is the base of the E-region. As already discussed, the velocity rises so rapidly in this region that the velocity profile is not greatly affected. Above the E-region

peak the velocity decreases with height, so that the corrected values given in Figure 1 will be smaller than the true velocities if the echoes are from below the $\mu = 0$ level. Hence these corrected velocities represent lower limits but the correction term, $(\overline{v}_T - v_I)\mu_I$, is likely to be small since both components of this term will generally be small. Of course ionosonde drift measurements at high-latitudes are dominated by reflections at or near the $\mu = 0$ level rather than by partial reflections or backscatter reflections.

As a final remark, in one respect there is a correspondence between drift velocity and virtual height in that each provides a measure of the integral effect of the ionosphere rather than an exact measurement of the property at the reflection point. However the results of this study show that, unlike virtual height, the drift velocity is a good estimate of the velocity at the reflection height in all but extreme situations. To emphasise this point, consider how well the drift velocity profile reproduced the shape of the velocity profile derived by inversion. This is in stark contrast to virtual height which, when plotted as a function of frequency, usually exhibits extreme cusps which vastly exaggerate the real height variation of plasma frequency.

14.4 Conclusions

The accuracy of ionosonde drift-velocity measurements has been examined for the situation in which the ionospheric velocity varies with height. This has been done by developing a simple inversion technique and applying it to an average drift velocity profile obtained at Casey, Antarctica. The results show that the ionosonde drift velocity technique provides reasonable results in the lower E-region and F-region. A narrow velocity peak near 110 km caused the drift velocity technique to underestimate the velocity at this level by as much as 20%. However the discrepancy dropped quickly at higher heights, falling below 7.5% at 150 km. The results support the use, at each measurement epoch, of a single estimate of the F-region convection drift at a high-latitude station such as Casey.

As with other measurements of ionospheric parameters, drift velocity profile measurements on a more rapid time scale will show much more variability than the average profile shown here. The results of this study suggest that such results could be validly interpreted provided it is recognised that such profiles may represent a smoothing of the true velocity profile in regions of rapid variation of velocity with height. While the inversion procedure developed here is simplified and appropriate for the average profile presented, it does provide a reasonable first estimate of the true velocity profile. A more rigorous approach will probably be required for 'instantaneous' velocity profile measurements as it will probably be necessary to consider the specific angles-of-arrival of echoes more explicitly.

14.5 Acknowlegements

This work has been supported by the Australian Research Council and the Australian Antarctic Science Advisory Committee. We thank J.A. Bennett for useful discussions and all Casey expeditioners who contributed to the maintenance and operation of the DPS-4.

References

Bennett, J.A. (1968). The ray theory of Doppler frequency shifts. *Australian Journal of Physics*, 21: 259–272.

Bennett, J.A. and Dyson, P.L. (1993). On Pfister's theorem and its generalization. *Journal of Atmospheric and Terrestrial Physics*, 55: 1307–1310.

Dyson, P.L. (1975). Relationships between the rate of change of phase path (Doppler shift) and angle of arrival. *Journal of Atmospheric and Terrestrial Physics*, 37: 1151–1154.

Dyson, P.L. and Benson, R.F. (1978). Topside sounder observations of equatorial bubbles. *Geophysical Research Letters*, 5: 795–798.

Gamache, R.R., Reinisch, B.W. and Kersey, W.T. (1992). ARTIST electron-density profile algorithm. *Scientific Report Number 468*. University of Massachusetts Lowell Center for Atmospheric Research.

Grant, I.F., MacDougall, J.W., Ruohoniemi, J.M., Bristow, W.A., Sofko, G.J., Koehler, J.A., Danskin, D. and Andre, D. (1995). Comparison of plasma flow velocities determined by ionosonde Doppler drift technique, SuperDARN radars, and patch motion. *Radio Science*, 30: 1537–1549.

Haines, D.M. and Reinisch, B.W. (1993). Digisonde portable sounder (DPS) series system. *Technical Manual, Version 3*. University of Massachusetts Lowell Center for Atmospheric Research.

Parkinson, M.L., Monselesan, D.P., Smith, P.R., Dyson, P.L. and Morris, R.J. (1997). Digital ionosonde measurements of the height variation of drift velocity in the southern polar cap ionosphere: initial results. *Journal of Geophysical Research*, 102: 24075–24090.

Pfister, W. (1971). The wave-like nature of inhomogeneities in the E-region. *Journal of Atmospheric and Terrestrial Physics*, 33: 999–1025.

Scali, J.L., Reinisch, B.W., Heinselman, C.J. and Bullett, T.W. (1995). Coordinated digisonde and incoherent scatter radar F region drift measurements at Sondre Stromfjord. *Radio Science*, 30: 1481–1498.

Smith, P.R., Dyson, P.L., Monselesan, D.P. and Morris, R.J. (1998). Ionospheric convection at Casey, a southern polar cap station. *Journal of Geophysical Research*, 103: 2209–2218.

Titheridge, J.E. (1985). Ionogram analysis with the generalized program POLAN. *Report UAG-93*. World Data Center A for Solar Terrestrial Physics, NOAA, Boulder, Co, USA.

15. HF PROPAGATION VIA THE F3 LAYER

A.A. Arayne[1], P.L. Dyson[1] and J.A. Bennett[2]

[1] Department of Physics
La Trobe University
Bundoora Victoria 3083 Australia
(email: p.dyson@latrobe.edu.au)

[2] Department of Electrical and Computer Systems Engineering
Monash University
Clayton Victoria 3168 Australia

Abstract

An additional ionospheric layer, the F_3 layer, has recently been identified in vertical incidence ionogram records from the equatorial station, Fortaleza in Brazil. This paper examines the HF propagation characteristics of this layer in the context of the signatures that would appear on backscatter and oblique ionograms. It is found that the relatively weak F_3 layer reported at Fortaleza will not produce a pronounced signature on backscatter ionograms, because the F_3 layer produces mostly relatively weak echoes at the same ranges as much stronger F_2 propagated echoes. In contrast, on oblique ionograms, F_3 layer signatures are likely to be observed emerging from the high angle F_2 trace. Of course if the F_3 layer is pronounced enough then distinct echo traces occur on both backscatter and oblique ionograms.

15.1 Introduction

Recently Balan *et al.* (1997, 1998) have shown that an additional layer, called the F_3 layer, can form in the equatorial ionosphere. In the late morning-noon period the upward drift produced by the combined effects of the **E** x **B** drift and neutral wind causes the F_2 layer to drift upwards, forming the F_3 layer, while the usual photochemical and dynamical effects maintain an F_2 layer at lower altitudes. Balan *et al.* (1997) presented bottomside ionograms from Fortaleza showing the presence of an F_3 layer in an ionogram sequence, and later Balan *et al.* (1998) presented occurrence statistics based on Fortaleza ionograms.

The presence of an F_3 layer may affect HF propagation, so there are possible implications for communication and surveillance systems, such as the Jindalee Operational Radar Network (JORN). This paper presents the results of an initial theoretical investigation into the basic effects of an F_3 layer on HF propagation.

15.2 F_3 layer structure

The F_3 layer can be identified by the appearance of an extra cusp on ionograms obtained by Vertical Incidence Sounders (VIS), as shown in the sequence of Fortaleza VIS ionograms presented by Balan *et al.* (1997) (Figure 1). However the identification may not be straightforward because travelling ionospheric disturbances (TID's) and the $F_{1.5}$ layer (e.g. Rishbeth and Garriot, 1969) produce similar cusps. The ionograms in Figure 1 have been scaled and inverted using the POLAN method (Titheridge, 1985) to obtain vertical electron concentration profiles. The profile at 1130 LT is shown in Figure 2 and it is apparent that the F_3 layer is quite weak, with f_0F_3 being just slightly greater than f_0F_2.

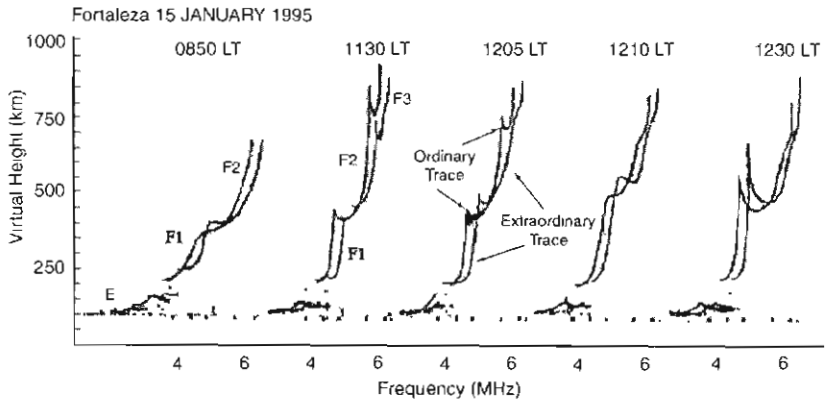


Figure 1. *Sequence of VIS ionograms from Fortaleza (3°S, 38°W), showing the presence of an F_3 layer (after Balan et al., 1997).*

Balan *et al.* (1997) modelled the F_3 layer using the Sheffield University Plasmasphere-Ionosphere Model (SUPIM). They considered sunspot maximum conditions (F10.7 = 178) rather than the more moderate level of sunspot activity that prevailed at the time of the Fortaleza F_3 layer observations they presented (F10.7 = 78). Their modelling predicted the formation of an F_3 layer about an hour before noon and lasting several hours. However f_0F_3 was predicted to be smaller than f_0F_2 at times and a ground-based VIS will only detect the F_3 layer when $f_0F_3 > f_0F_2$. At Fortaleza this occurs for typically three hours, although a duration as short as 15 minutes and as long as six hours have been observed (Balan *et al.*, 1998).

The vertical electron concentration profile from the model results of Balan *et al.* (1997) is also plotted in Figure 2. It is taken from Balan *et al.* (1997) where the profile does not extend below 200 km, so in Figure 2 it has been extended using appropriate IRI parameters. The difference

between the profile obtained from the ionogram inversion and the model profile is quite stark. The electron concentrations and heights of the F_2 and F_3 layers are much greater in the model case, which corresponds to higher solar activity.

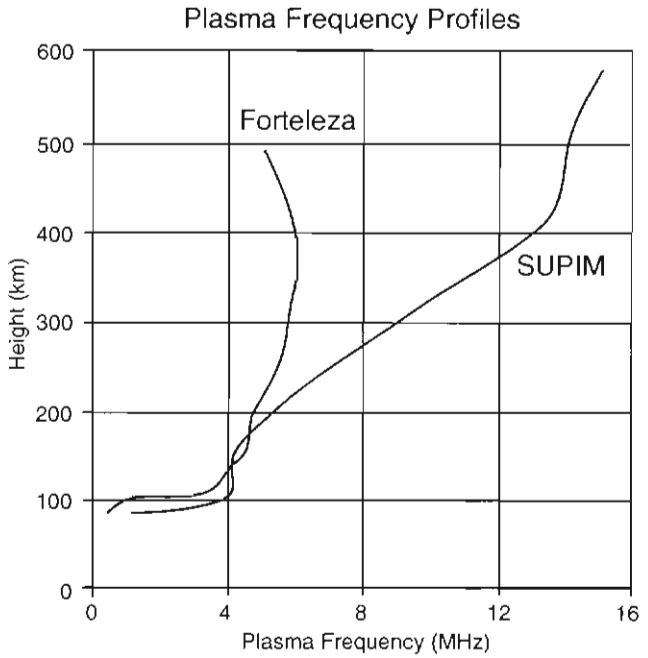


Figure 2. *Vertical ionospheric profiles at 1130 UT. Fortaleza profile obtained by inversion of corresponding ionogram in Figure 1. SUPIM profile adapted from Balan et al. (1997).*

15.3 Ionospheric model and ray tracing

Oblique Incidence Sounders (OIS) and Back-Scatter Sounders (BSS) are used in systems such as the Jindalee Operational Radar Network (JORN) to monitor HF propagation conditions in real time, and provide information for frequency management and target location, normally referred to as Coordinate Registration (CR). As explained above, the F_3 layer produces an identifiable signature on VIS ionograms and the aim here is to investigate what, if any, characteristic features will occur on OIS and BSS ionograms when F_3 layer propagation occurs. This has been pursued by specifying models of the ionosphere, including an F_3 layer, and using ray tracing to determine the corresponding OIS and BSS ionograms.

Our knowledge of F_3 layer behaviour is still very limited. Balan *et al.* (1998) have completed a study of its occurrence at Fortaleza but there has been no study of its spatial extent. In fact one of the motivations of this study is to examine whether OIS's and BSS's could readily contribute to morphological studies of the F_3 layer. In the absence of any real

knowledge of the spatial extent of the F_3 layer, we confine our study to the effects of the vertical profile on oblique propagation. This is not as restrictive as might be first thought, since, if the mid-point vertical profile is known, horizontal linear gradients have no first order effect on group path, the parameter displayed on VIS and BSS ionograms. Hence unless the F_3 layer produces highly non-linear horizontal structure, the use of a single vertical profile will be valid when applied to the mid-point of propagation paths. Hence it is applicable to OIS ionograms. In the case of BSS ionograms, no single mid-point profile is applicable, nevertheless, the use of a single vertical profile will highlight effects arising from the F_3 layer vertical structure.

We have synthesised OIS and BSS ionograms using the analytical ray tracing approach of Dyson and Bennett (1988), Chen *et al.* (1990) and Bennett *et al.* (1991). Croft and Hoogasion (1968) showed that describing the ionosphere by a quasi-parabolic vertical profile gave rise to analytical solutions for ray parameters for propagation over a spherical Earth. Dyson and Bennett (1988) and Chen *et al.* (1990) extended this concept to the use of multi-quasi-parabolic segments to describe vertical profiles of arbitrary structure. In this approach the effect of the Earth's magnetic field on the propagation is generally ignored although it can be included (Chen *et al.*, 1992). The main effect of the Earth's magnetic field it to cause birefringence, so that on ionograms two echo traces are observed, corresponding to ordinary and extraordinary mode propagation. However these two traces generally exhibit similar characteristics so we can expect the F_3 layer to produce the same major effects on both traces. Hence we can safely ignore the Earth's magnetic field and hence birefringence.

OIS and BSS ionograms have been synthesised using the vertical profiles shown in Figure 2. While the profiles appear quite different, general features of the F_3 layer propagation are similar. The reason is related to the fact that, when the Earth's magnetic field is ignored, propagation effects scale as f/f_c where f is the propagation frequency and f_c is the maximum ionospheric plasma frequency (normally f_0F_2, but in these cases f_0F_3). The two profiles in Figure 1 are actually very similar in the F_2 and F_3 regions when they are normalised in this way (i.e. with $f_c = f_0F_3$), and the heights, h, are also normalised using h/h_{max}, where h_{max} is the height of the F_3 layer peak. In both cases f_0F_2/f_0F_3 is close to 0.90. Consequently we only show results for one case, viz., the SUPIM model profile of Figure 2.

15.4 F_3 layer signatures on back-scatter ionograms

For the ionospheric profiles shown in Figure 1, BSS ionograms have been synthesised using the general purpose analytical ray tracing package, QPSHEL (Dyson *et al.*, 1994). Since we are interested purely in the effect of the ionosphere on the propagation, and BSS system losses have been ignored and isotropic antennas have been assumed. The synthesised BSS ionogram for the SUPIM profile is shown in Figure 3.

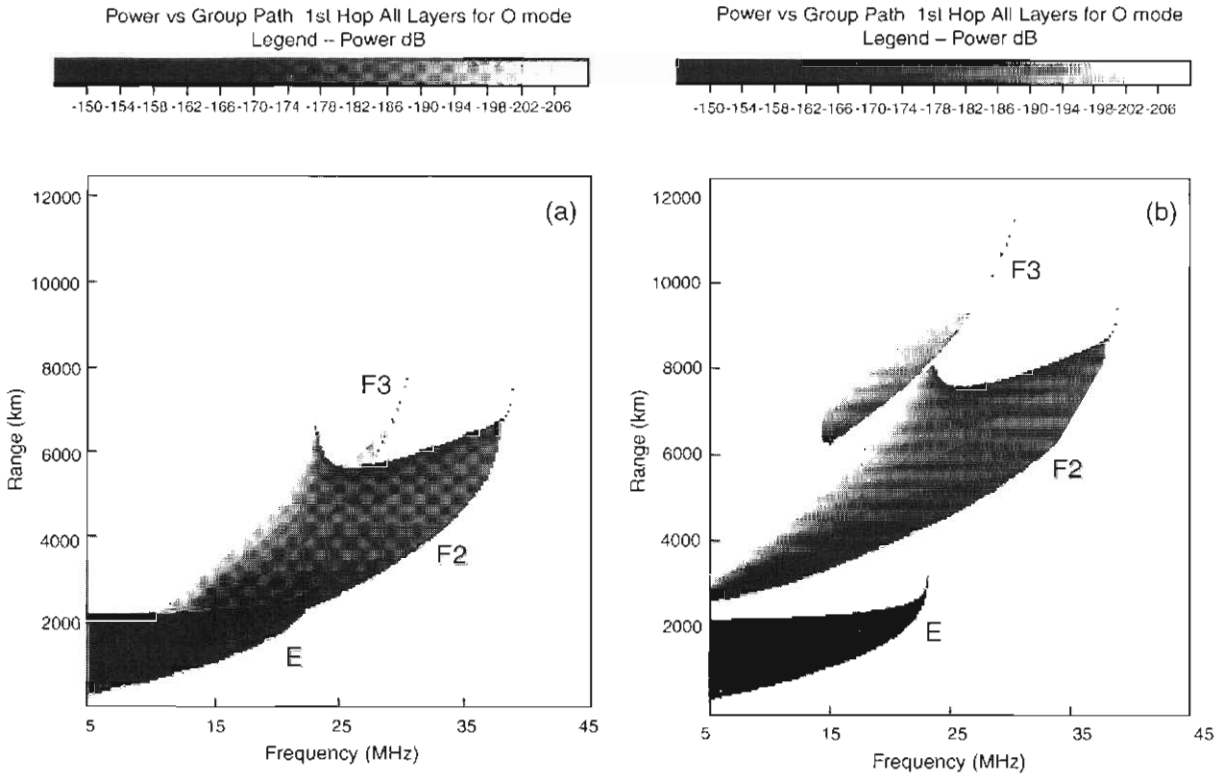


Figure 3. *(a) Backscatter ionogram synthesised using the SUPIM profile from Figure 2. (b) Layer traces displaced in group range.*

It is apparent that the F_3 layer produces a relatively weak echo trace which is largely hidden behind the E and F_2 layer traces. This occurs because f_0F_3 is not much greater than f_0F_2 but at a greater altitude.

In a spherically stratified ionosphere Snell's Law become Bougere's Law and may be written as

$$r\,\mu\cos\beta = \text{constant}$$

where r is the radial distance,
μ is the refractive index
and β is the elevation angle at r.

$$r\,\mu\cos\beta = r_E\cos\beta_E$$

where r_E is the radius of the Earth
and β_E is the elevation angle at the Earth's surface.

At reflection a ray is horizontal, so

$$r_R\,\mu_R = r_E\cos\beta_E$$
or $$(\mu_R)^2 = (r_E\cos\beta_E / r_R)^2$$

Now, $$(\mu_R)^2 = 1 - (f_n/f_R)^2$$

where f_n is the plasma frequency
and f_R is the radio propagation frequency reflected at r_R.

Hence f_{RM}, the maximum value of f_R corresponds to the maximum value of l_R, which occurs when $\cos\beta_E = 1$, i.e. when $\beta_E = 0$.

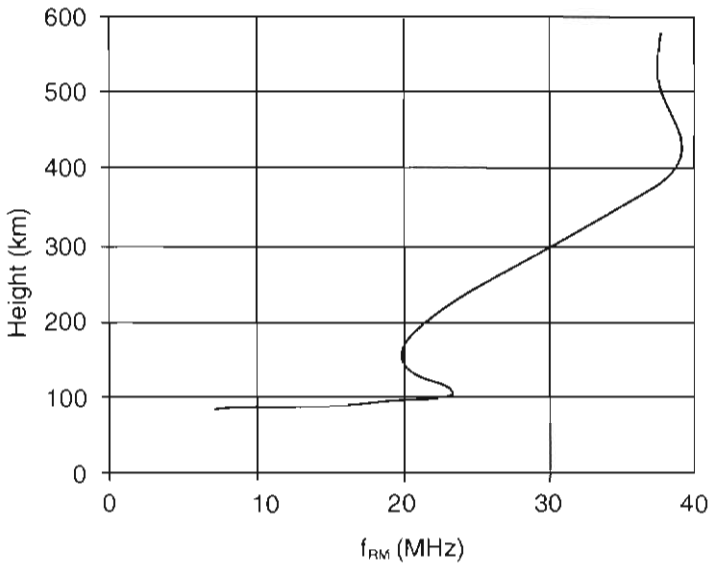


Figure 4. *Maximum frequency that can be reflected by the plasma frequency at a particular height.*

Figure 4 shows the variation of f_{RM} for the SUPIM profile. It is apparent that, consistent with the synthesised BSS shown in Figure 3, the F_2 layer can reflect higher frequencies than the F_3 layer. Of course the actual maximum frequencies reflected by each layer will be less than those shown in Figure 4 if the $\beta_E = 0$ and other low angle rays are reflected by lower layers. This is certainly true for the F_3 layer where the maximum frequency propagated by the layer is less than 31 MHz.

By way of contrast, a model vertical profile with a much more pronounced F_3 layer (Figure 5(a)) has also been used to synthesise a BSS ionogram. In this case a pronounced F_3 layer trace is produced (Figure 6) as would be expected from the plot of f_{RM} versus height (Figure 5(b)).

It is worth noting that because of the high altitude of the F_2 and F_3 layer peaks, both model ionospheres support single hop propagation over distances in excess of 6000 km.

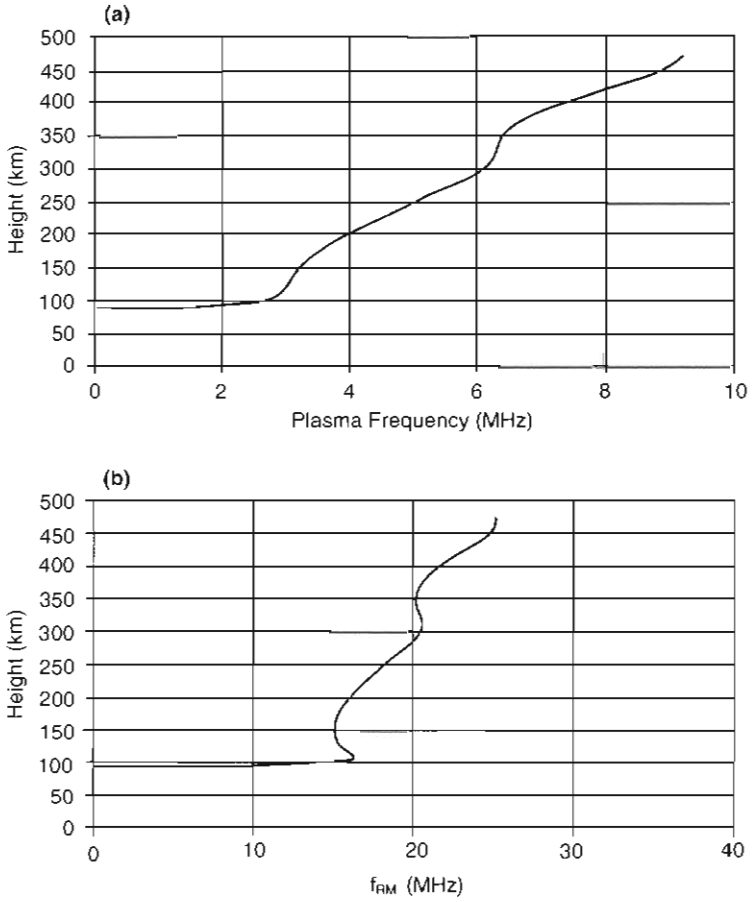


Figure 5. *(a) Vertical ionospheric profile. (b) Maximum frequency that can be reflected by the plasma frequency at a particular height.*

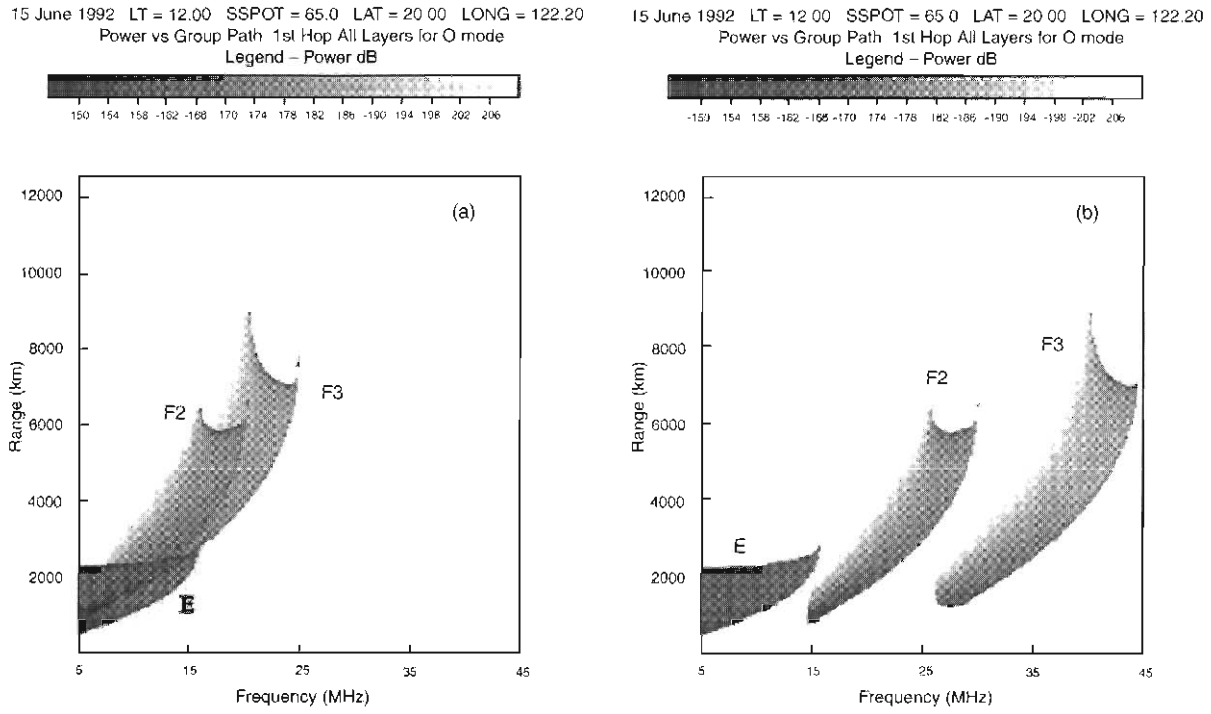


Figure 6. *(a) Backscatter ionogram synthesised using the profile from Figure 5(a). (b) Layer traces displaced in frequency.*

15.5 The F_3 layer OIS signature

Figure 7 shows the oblique ionogram for a 1250 km path synthesised using the SUPIM profile. There is a distinct F_3 layer trace emerging from the high angle F_2 layer trace. Ionograms synthesised for other paths show a similar effect. This trace signature has been observed on OIS's operating over paths extending north of Australia (Lynn, private communication). Such observations may be indications of the presence of an F_3 layer although this signature, of an additional layer trace emerging from the high angle F_2 layer trace, can also be produced by travelling ionospheric disturbances. From comparison of Figure 7 and Figure 3 it is evident that the presence of a weak F_3 layer is more likely to be apparent on OIS ionograms than on BSS records.

The equivalent synthesised oblique ionogram for the more pronounced F_3 layer profile of Figure 5, is shown in Figure 8. In this case the F_3 layer is more pronounced, with three very obvious layer traces (E, F_2 and F_3). In practice, the layers might be mistakenly identified as E, F_1 and F_2 or E, $F_{1.5}$ and F_2. In a practical case it would be important to examine the time sequence of OIS ionograms, as is required for positive identification of F_3 layer from VIS ionograms. Comparison of Figure 8 and Figure 6 shows that in this case the F_3 layer is clearly evident on both OIS and BSS ionograms.

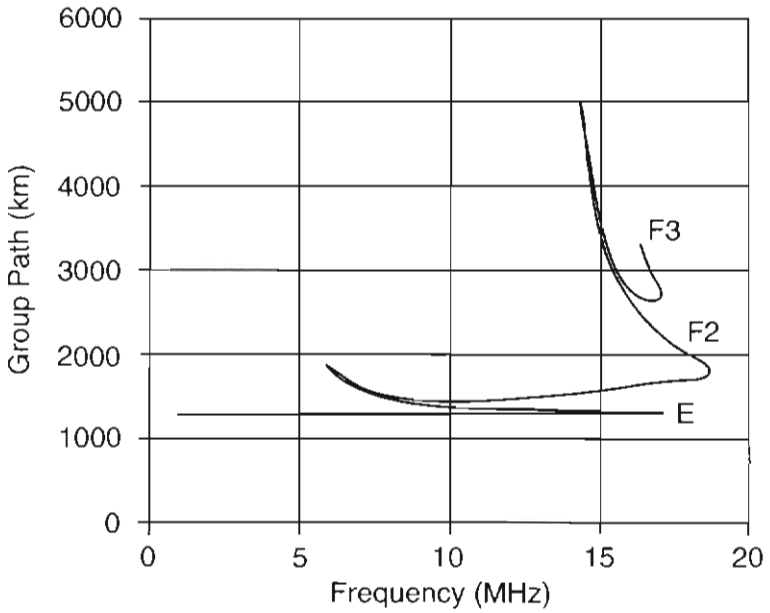


Figure 7. *Oblique ionogram for SUPIM profile in Figure 2.*

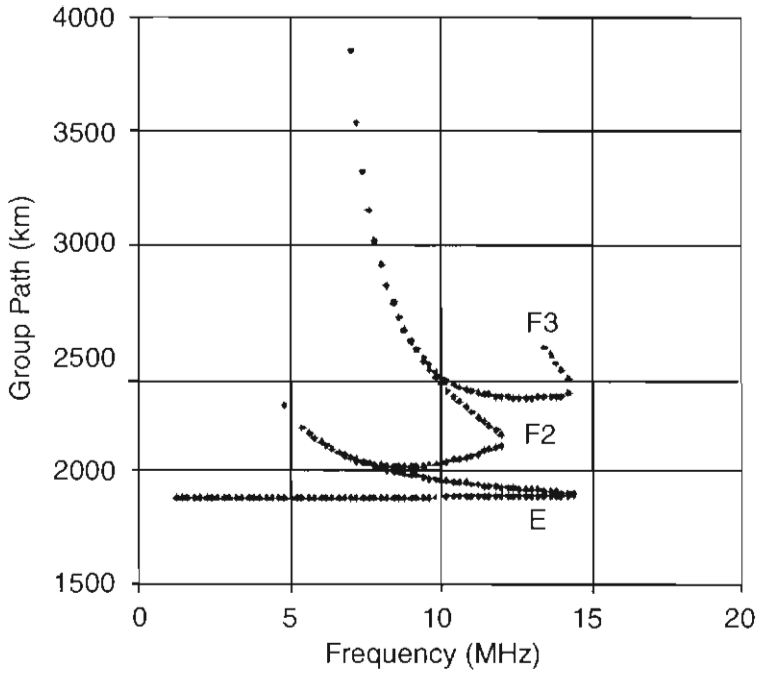


Figure 8. *Oblique ionogram for ionospheric profile in Figure 5.*

15.6 Conclusions

The expected effects of the F_3 layer on OIS and BSS ionograms has been investigated by synthesising ionograms using model vertical profiles of the ionosphere. Two models have been used. One contains a weak F_3 layer, consistent with reported VIS observations and corresponding to an F_3 layer model developed by Balan *et al.* (1997). The second model contains a pronounced F_3 layer.

If the F_3 layer is weak, then it is unlikely to be observed by BSS sounders because strong echoes propagated by lower layers (primarily the F_2 layer) will be superimposed on the weak F_3 layer echo trace. However the F_3 layer will be apparent on OIS ionograms as a trace emerging from the high angle F_2 layer trace and this is observed.

If the F_3 layer becomes pronounced enough so that it supports oblique propagation over long distances at higher frequencies than the F_2 layer, then very distinct traces may occur on both BSS and OIS ionograms. These traces exhibit the same distinct features displayed by traces caused by the other ionospheric layers. In principle the presence of an F_3 layer is then readily identifiable although in practice care will be required to avoid misidentification as an F_2 layer, followed by consequential misidentification of other layers.

15.7 Acknowlegements

This work has been supported by the Australian Research Council. We thank Dr Ken Lynn for helpful discussions and access to data prior to its publication elsewhere.

References

Balan, N., Bailey, G.J., Abdu, M.A., Oyama, K.I., Richards, P.G., MacDougall, J. and Batista, I.S. (1997). Equatorial plasma fountain and its effects over three locations: evidence for an additional layer, the F_3 layer. *Journal of Geophysical Research*, 101: 2047–2056.

Balan, N., Batista, I.S., Abdu, M.A., MacDougall, J. and Bailey, G.J. (1998). Physical mechanism and statistics of occurrence of an additional layer in the equatorial ionosphere. *Journal of Geophysical Research*, 102: 29169–29181.

Bennett, J.A., Chen, J. and Dyson, P.L. (1991). Analytic ray tracing for the study of HF magneto-ionic radio propagation in the ionosphere. *Applied Computational Electromagnetics Society Journal*, 6: 192–210.

Chen, J., Bennett, J.A. and Dyson, P.L. (1990). Automatic fitting of quasi-parabolic segments to ionospheric profiles with application to ground range estimation for single-station location. *Journal of Atmospheric and Terrestrial Physics*, 52: 277–288.

Chen, J., Bennett, J.A. and Dyson, P.L. (1992). Synthesis of oblique ionograms from vertical quasi-parabolic segment models of the ionosphere. *Journal of Atmospheric and Terrestrial Physics*, 54: 323–331.

Croft, T.A. and Hoogasian, H. (1968). Exact ray calculations in a quasi-parabolic ionosphere. *Radio Science*, 2: 69–74.

Dyson, P.L. and Bennett, J.A. (1998). A model of the vertical distribution of the electron concentration in the ionosphere and its application to oblique propagation studies. *Journal of Atmospheric and Terrestrial Physics*, 50: 251–262.

Dyson, P.L., Bennett, J.A. and Chen, J. (1994). Analytic ray-tracing program–QPSHEL93. *Contract Report*. School of Physics, La Trobe University. p. 64.

Rishbeth, H. and Garriot, K.O. (1969). *Introduction to Ionospheric Physics*. Academic Press, New York.

Titheridge, J.E. (1985). Ionogram analysis with the generalised program POLAN. *Report UAF-93*. World Data Centre A for Terrestrial Physics, NOAA, Boulder, Co, USA.

16. OBSERVATIONS OF POLAR CAP ARC DRIFT MOTION FROM SCOTT BASE

I.M. Wright, B.J. Fraser and F.W. Menk

Cooperative Research Centre for Satellite Systems
Department of Physics
University of Newcastle
NSW 2308 Australia
(email: bjfph@cc.newcastle.edu.au)

Abstract

Apparent drift motion of polar cap arcs may be indicative of plasma convection processes in the ionospheric plasma. The dynamics of a number of individual polar cap events observed by a dual wavelength all-sky imager at Scott Base have been analysed in detail. The imager, which records optical emissions at 630 nm and 428.7 nm has been operated at Scott Base, (79.96°S, 327.61°E magnetic), during the winters of 1996 to 1998. The events presented have been selected to show representative examples of arcs observed primarily in the dayside. Of interest is the speed and direction of the polar cap arc drift motion with respect to solar wind and interplanetary magnetic field conditions. Events were observed between 10:00 and 19:00 local time (mid-morning to dusk), where drift velocities varied between 60 m s^{-1} away from the invariant pole and 1000 m s^{-1} polewards. Another example shows a drift velocity of 300 m s^{-1} perpendicular to the pole direction. Each of the examples show different velocities and drift direction, for differing solar wind condition, at different times. The small number of images in the initial data set needs to be expanded into a three year data set, and include data from other instruments, in order to fully understand the correlations of polar cap arc drifts with solar wind/interplanetary magnetic field (IMF) conditions.

16.1 Introduction

The observation of optical auroral forms that occur in the polar cap has led to the identification, and naming, of a wide range of emissions including: polar cap arcs, extremely high-latitude auroras, sun aligned arcs, transpolar arcs, horse collar auroras, and theta aurora. It is difficult to interrelate these as they have been defined by various authors under a wide range of conditions and over a period of observations that began in 1907. In a recent review paper, Zhu *et al.* (1997) use the term polar cap arc to differentiate aurora seen in the polar cap or poleward of the auroral oval, from auroral oval emissions. Unlike auroral zone arcs these are seen more during northward IMF and quiet magnetic conditions. Early observations, made visually, and by film based all-sky

camera, identified properties of polar cap arcs including the Sun-Earth orientation, transverse motions, the occurrence only at high-latitude, and the negative correlation with magnetic activity.

The launching of low Earth orbit satellites in the 1970's, for example the DMSP series, enabled particle measurements to be related to optical emissions (Hardy *et al.*, 1982). However, the orbits limited these studies to relatively local regions, and spin scanning resulted in coarse resolution in time and position. UV imaging and the high altitude elliptical orbit of DE 1 in the 1980's led to the first view of the theta aurora spanning the entire polar cap, a phenomenon that ISIS 2 and DMSP could not see. The DE-1 and 2 satellites found that polar cap arc precipitation had plasma sheet or plasma sheet boundary layer (PSBL) characteristics suggesting these arcs occur on closed field lines (Peterson and Shelley, 1984; Frank, 1988). This was supported by the observed conjugacy of transpolar arcs (Craven *et al.*, 1991). However, others, (e.g. Hardy *et al.*, 1982; Gussenhoven and Mullen, 1989), have seen polar rain and relativistic electrons near polar arcs suggesting a source on open field lines. It is still not clear whether polar cap arcs occur on open or closed field lines, or whether the same phenomenon can arise under either of these conditions. (The review by Zhu *et al.* (1997) contains comprehensive references.)

Continued observations of high-latitude optical emissions are needed to resolve questions about their relationship to open/closed field lines, and equally importantly, to the ionosphere convection that accompanies them. This study uses images of polar cap aurora acquired at Scott Base, Antarctica to investigate the scale, the motions, and dynamics of these phenomena to examine these questions. In particular the behaviour of small scale, relative to the polar cap, arcs of less than 1000 km length are noted.

16.2. Scott base imager project

16.2.1 Imager site

An all-sky digital imager has been operating at Scott Base since 1996 to record optical phenomena during the darkness of the Antarctic winter. The first season of operation produced 28 days of data when aurora was visible for some time during the 24 UT hours, over the 1996 Austral winter. In 1997, the second season, 22 days of useful data were obtained. All data from the first two seasons were recorded during quiet magnetic conditions, $K_p \leq 2$. The presence of cloud, or interference from Sun or Moon light, prevented collection of images at other times. For 1998, there were 36 days of good data recorded, some during more active conditions. Out of the data from the three operation periods, on 38 occasions the drift motion of detected arcs have been selected for further study. The intensity and structure of arcs can vary during their lifetime. The average lifetime is some tens of minutes, but they occasionally exist for longer than one hour.

New Zealand's Scott Base station is at Pram Point on Ross Island in the Ross Sea, and the USA station, Mc Murdo, is a few kilometres away at Discovery Point. Both New Zealand and the USA each have a permanent laboratory at Arrival Heights which is about 60 metres above sea level, and four kilometres inland from the coast. Arrival Heights is located at 77.82°S, 166.67°E geographic and is 79.96°S, 327.61°E CGM magnetic. This geographic location experiences total darkness at local noon for several weeks as the Sun is below the horizon during winter. These conditions allow the study of dayside emissions that normally are invisible.

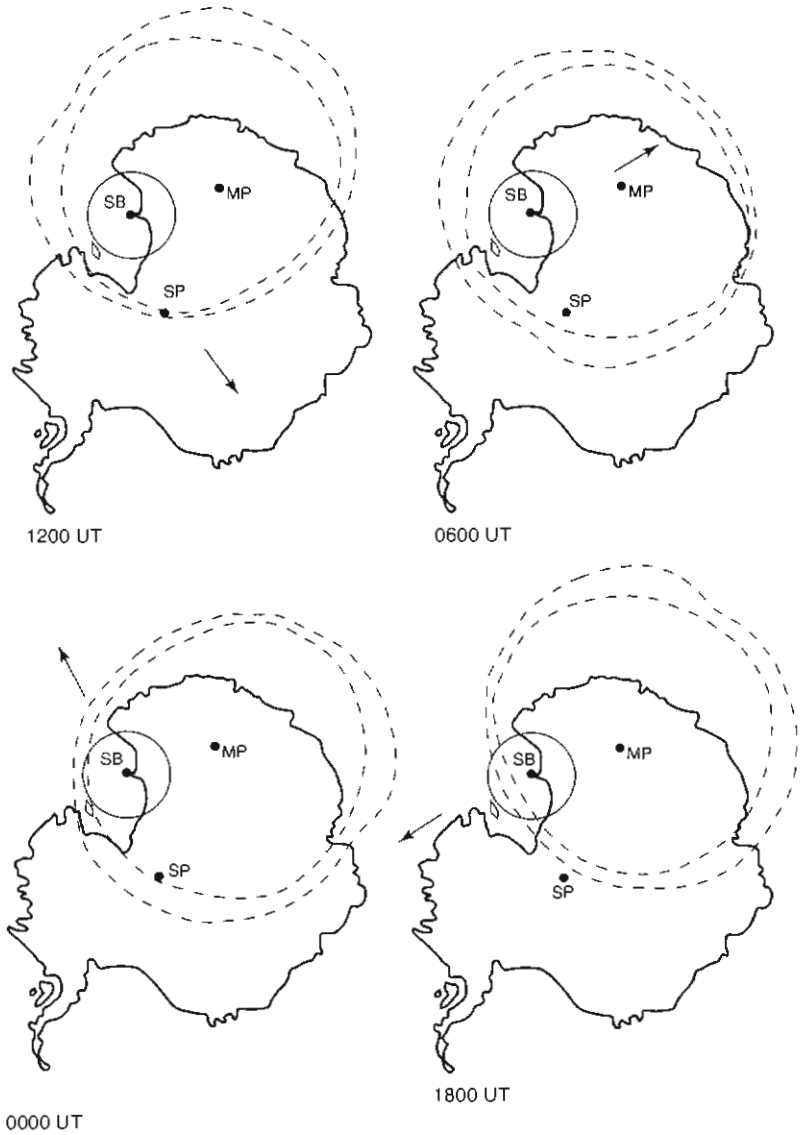


Figure 1. *Imager field-of-view looking upwards from Scott Base.*

Figure 1, superimposes the quiet time (K_p ~ 2) auroral oval of Feldstein and Starkov (1967) for each 6 hour interval over the map of Antarctica. It shows that Scott Base (SB) is completely within the polar cap, except for the hours around 1800 UT, when the quiet time auroral oval may also enter the field-of-view. The circle, centred on Scott Base, represents a 1000 km diameter field-of-view (see Section 16.2.3), and the auroral oval is shown rotating around the fixed imager location. The direction to the Sun is indicated by the arrows. This figure shows the upwards looking imager view, and the superimposed map of Antarctica is reversed left to right. In the images, the direction to the geomagnetic pole (MP) is to the upper left at 45°. The geographic pole is labelled 'SP'. At Scott Base, local solar noon is at 0056 UT.

16.2.2 Dual wavelength all-sky imager

Observations are recorded as two side by side circular horizon-to-horizon images of the sky. The left image contains 630 nm wavelength information and the right image is of 427.8 nm. There are examples presented in the later figures. The all-sky digital imager is housed in the Arrival Heights laboratory using a light trap below a glass dome in the roof.

Room temperature air from the laboratory is circulated through the light trap and dome to minimise icing inside the dome. Figure 2 shows the schematic arrangement of the imager. It was manufactured by Keo Consultants of Boston and is similar to units used by the USA at South Pole station and the US PENGUIn research project. Keo Consultants also supplied a different version of the instrument to the Australian Antarctic Division. A Canon 16 mm focal length 'fish eye' lens creates a horizon-to-horizon image of the sky. This image is collimated so essentially parallel rays are passed through a dual band-pass interference filter that only passes 630 nm and 428.7 nm wavelengths to a beam splitting prism. The prism produces the two side by side images of the whole sky, one image has the 427.8 nm wavelength removed by another interference filter, and the other has the 630 nm information removed.

These side by side images, are focussed onto an image intensifier. Keo Consultants documentation states the system detects minimum emissions of 50 Rayleigh. An intensifier is needed to detect optical emissions of this level. This intensifier was developed for night-time military use and has ramifications that will be discussed later. The output intensifier image is scanned by a CCD television camera. The camera includes a power supply for the intensifier, the TV camera, a rough calibration light, and a mechanical shutter below the objective lens that is interlocked with a light detector. The shutter and light dependant resistor protect the image intensifier from damage that would result if an image of the Sun was allowed to fall on the input.

Under normal operation the all-sky imager will take a pair of exposures every 30 seconds. One exposure of 0.9 s duration, and the other 0.3 s

are achieved by allowing the CCD to integrate for these periods by suspending the normal television scanning at 25 frames per second. These automatic gain controlled images are 8-bit digitised, compressed, and then written to magneto-optical disk. In this paper we concentrate on the 0.9 s exposures which provide more detailed images than the 0.3 s exposures. The operation of the system is software controlled from a dedicated PC, with a video monitor for image display.

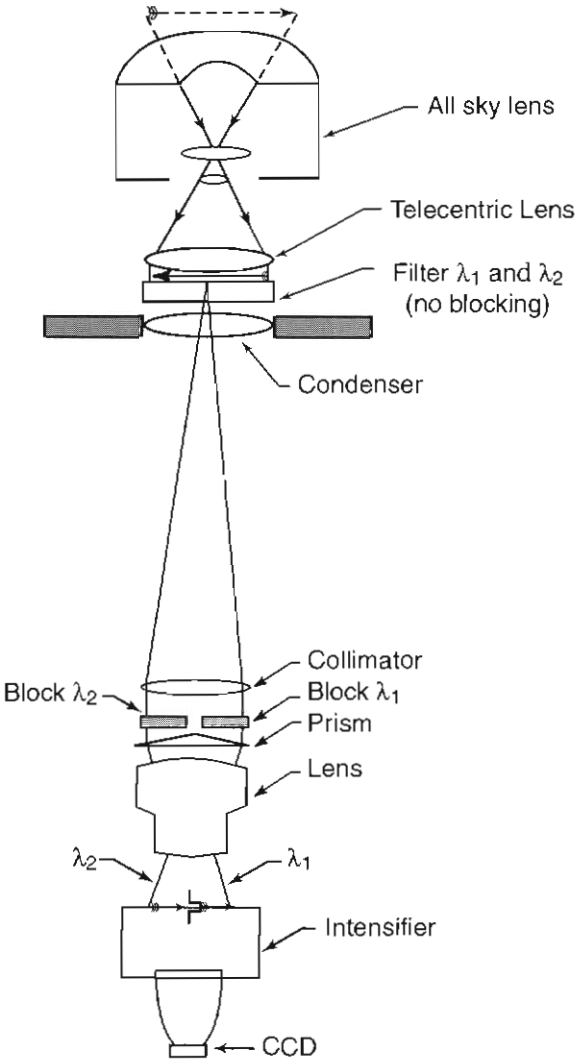


Figure 2. *All-sky imager schematic diagram.*

16.2.3 Data characteristics

Auroral properties observed from the ground include; apparent extent, apparent velocity, and apparent intensity. From a single observation point, assumptions need to be made to estimate actual values for these parameters. These assumptions depend on the performance of the of the observing device, and experience based on observations and results from previous work.

The oxygen-electron reaction that produces the red 630 nm emission occurs at altitudes from 200 km to as high as 450 km, depending on atmospheric conditions such as the concentration of oxygen (Carlson and Egeland, 1995). The nitrogen–electron reaction responsible for the blue 427.8 nm emission only occurs lower in the ionosphere, at about 120 km altitude. Estimates of size and movement of auroral events need to take the altitude into consideration.

An arbitrary assumption we make is that the red image occurs at an altitude of 250 ±50 km, a figure used by other workers (e.g. Sandholt *et al.*, 1998), which is twice the assumed altitude of the blue image. A 1000 km diameter field-of-view used in this study corresponds to the 120 km altitude maximum field-of-view of the lens. The geometry of the fish-eye lens compresses the image to the edges of the field so that the central 250 km diameter of sky occupies half the image. The resolution of the 256 x 256 pixel system is very coarse at the edge of the field-of-view and only very large-scale features can be discerned away from the centre of the images. Small scale arcs, relative to the scale of the polar cap, are resolved in the centre 250 km area. The behaviour of narrow arcs of less than 1000 km in length are studied here.

The recorded image is dependent on the instantaneous response of a photon cascade device (the image intensifier) and consequently there is an inherent uncertainty in the level of the recorded signal. The effect of the image intensifier is shown in Figure 3. This shows the response of the imager system during tests made to back calibrate the first season's data on the laboratory bench. A low light intensity test target using light emitting diodes (LED's) was constructed. This consisted of a grey square mounted in a matt black light proof box. The camera could be operated seeing only this square, filling the central third of the field-of-view, with all extraneous light excluded. The voltage fed LED light source could be varied to produce images with the same intensities as natural signals found in observed arcs. The images recorded from the test box consist of a square in the centre of each of the red and blue side by side images. Allowing for the uneven illumination in the test box, a horizontal scan of intensities across the centre line of this test image showed an increase and then a decrease at the target square boundaries in each of the red and blue images, with a reasonable plateau between.

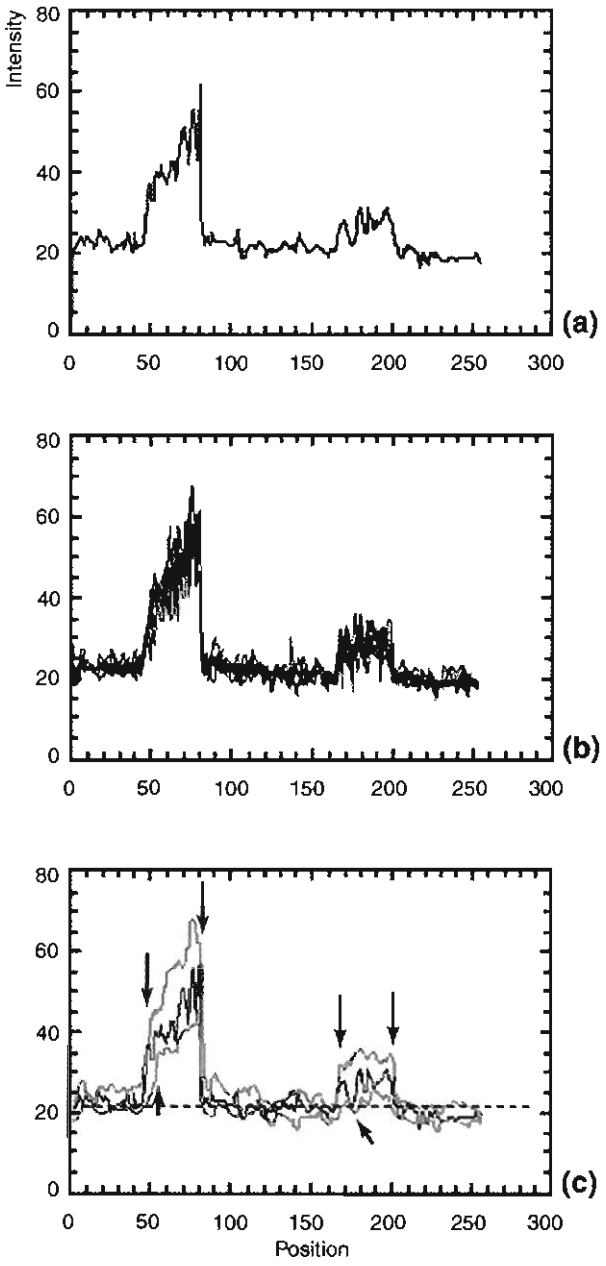


Figure 3. *Imager system responses to test target; (a) response for one frame and (b) stacked responses from six frames (c) upper and lower limits from (b).*

The plot in Figure 3(a) shows the response across line 128 (out of 256 lines) in one 0.9 s image recorded from the test target. The y-axis values are arbitrary pixel intensities which vary between 0 and 255, as determined by the intensifier gain and the compression software. In this instance the maximum is about 65, comparable with intensities from many auroral events. The x-axis is the pixel location number across the image. The edges of the square in the red image are seen at pixel 48 and 90, and in the blue image at about pixel 153 and 200 in Figure 3(a). Ideally, all test images recorded under identical conditions should be the same. In Figure 3(b), the response for another five images (0.9 s) recorded successively every two seconds due to the variability of the intensifier, the only boundary edge that remains well defined is at pixel 90.

In Figure 3(c) the range of the six traces from Figure 3(b) is shown, with arrows marking points of interest. The edges remain reasonably well defined in the red and blue images, provided that the response was above the noise level. This is indicated by the vertical arrows. The approximate background noise (intensity level 22 ±3 in the field data, and in the tests) is indicated by the horizontal dashed line. The oblique arrow shows where the response in the blue image is sufficiently low to be lost in system noise even when a blue target is present.

It is seen that successive images of a constant target are not identical pixel by pixel, but can vary by about 25%. Averaged over time and area the variable response of the intensifier should, for a constant signal, provide a meaningful average response. However, for this application of an image intensifier a significant error must result when capturing single images of aurora that vary in both time and space. The recorded intensity could be anywhere within the 25% band. The recorded position of edges of arcs in the images has an error range of about 5%. By estimating the centre of an arc in the red images, and allowing for altitude, reliable velocity estimates can be made.

There are many examples in the field data of blue arcs apparently switching on or off in successive exposures, this is explained by the system's blue response, shown in Figure 3(c). As the result of the testing, it is apparent that absolute calibration of the system is not possible, and intensity ratios of instantaneous red/blue images are unreliable. Intensity ratios have not been used in this study.

16.3 Observations

The all-sky data recorded over the Antarctic winter has provided many images where the aurora has been well defined, and of sufficient intensity, to make reasonable deductions of spatial and temporal characteristics using the altitude assumptions mentioned above. The motion of polar cap arcs may display an aspect of overall ionospheric convection, or may be quite localised. Examples of polar cap auroras discussed here have been selected for events displaying a definite drift motion. The prevailing IMF conditions (projected from upstream) which indirectly drive ionospheric convection are also presented for these

events. These examples show that polar arc drift motion can be quite varied.

16.3.1 20 June 1996

On 20 June 1996 between 23:23:00 and 23:31:30 UT, as shown in Figure 4, a very bright, broad, roughly sun aligned arc formed from the SSE, where it was attached to the auroral oval, extending to the NNW. It moved poleward, (to the west) as a loop, and then collapsed to the south within six minutes. At the south-east, where this arc was connected to the auroral oval, it did not show much movement. The auroral oval, which should be seen in the lower left of the images was partly obscured by icing inside the dome. The thumbnail, Figure 4(a), shows the raw image of 23:27:00 UT, and the composite negative image. Figure 4(b), shows the positions of the arc every 30 s, moving from east to west for the period in the red image. The blue image (right side) shows positions of the arc between 23:25:00 and 23:28:30 UT. Here is an example of the blue image apparently switching on and off, as mentioned above. Assuming an altitude of 250 km for the red emissions, the westward velocity was 500 ±80 m s^{-1} over the centre of the field-of-view. At this time the direction to the Sun is just east of north, and is indicated by the arrowhead. Figure 4(c) shows the solar wind conditions from the WIND spacecraft Magnetic Field Instrument (MFI) data. The vertical line labelled '23:24 eq' relates to the conditions expected at the Earth at the time of the event, taking into consideration the prevailing solar wind

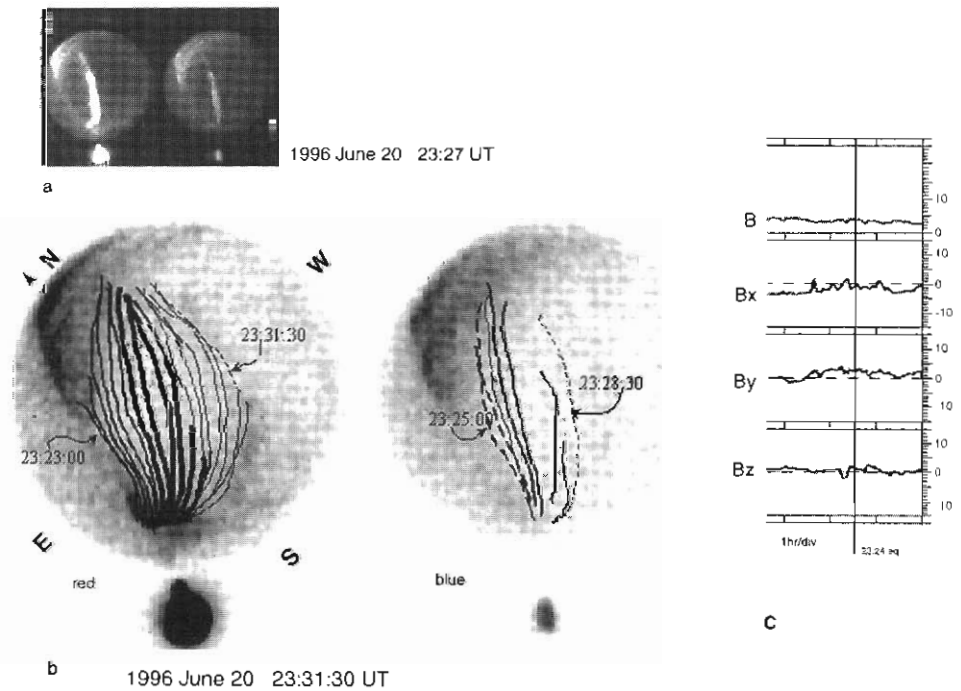


Figure 4. *20 June 1996 event, (a) raw data image, (b) arc positions (c), and WIND MFI data.*

velocity and travel time to the magnetopause. K_p = 1-, B_z was weakly positive, having been so mostly for the previous hours, B_y was also weakly positive and B_x at 23:24 eq was negative. The total magnetic field B was about 5 nT. There were two other similar arc events, with similar short lives, though not as intense, during the period 22:00 UT 20 June to 00:35 UT 21 June, around midday local time, with drift motion towards the invariant pole.

16.3.2 17 July 1996

The thumbnail, Figure 5(a), shows the raw image seen at 21:52 UT. This event also occurred just before local noon. A Sun aligned arc is visible in both red and blue images above dawn and cloud interference. The arc drift motion is eastwards away from the magnetic pole. The composite image, Figure 5(b) shows positions of this arc, which had a slower drift motion, every two minutes between 21:37 and 21:59 UT. At the end of this period the arc collapsed lengthwise to a vestige patch in the south, after experiencing eastward drift velocities of between 60 ±8 m s^{-1} and 200 ±25 m s^{-1}. The higher velocity was seen in the south. K_p = 2 at 21:00 UT. As seen in Figure 5(c), where the vertical line indicates the conditions projected to a time within this event, B_z = 0, and had been for the preceding hour. B_y, $B_x \geq$ -2 nT with B_{tot} = 5nT. For the drift direction of this arc B_y was negative while for the 20 June 1996 event described above B_y was positive.

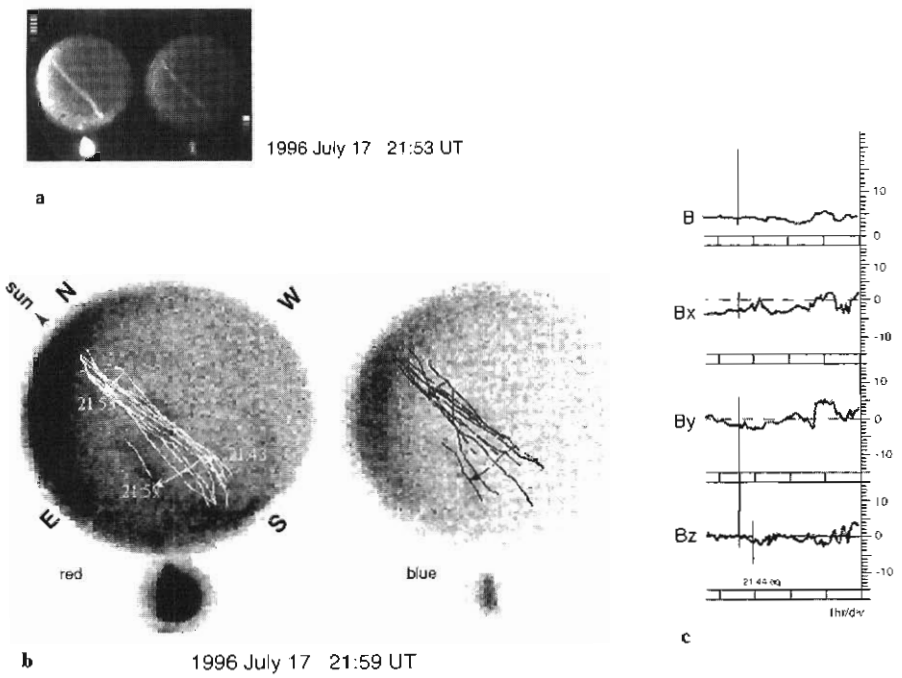


Figure 5. *17 July 1996. (a) raw data image; (b) arc positions; (c) WIND MFI data.*

16.3.3 22 July 1997

The raw image, Figure 6(a), is shown for 22:10:00 UT, one of the images used to make the composite image in Figure 6(b). This shows the locations of this arc every 30 s, in the red image, between 22:08:00 UT, the initial position, and 22:14:00 UT when it faded from view. Again the blue image is visible for less time than the red one. There is interference from the Sun at the top left of the images, and the Moon at the top right, and again there is some ice inside the dome on the lower left side of the images. This arc, another example before local noon, apparently rotated about its northern end. The velocity overhead was about 500 ±65 m s^{-1} while at the south end it could have exceeded 1000 ±125 m s^{-1}. Again in Figure 6(c) the MFI data with the prevailing conditions indicated shows the total B_{tot} field strength was of the order of 5 nT and had been stable for the previous hour or two, B_y = +2 nT, and $B_x \geq$ -2 nT this time. B_z was positive, about +2 nT, before the event, and although fluctuating, it remained mostly positive throughout the event. While the north end of this arc remained fixed, the rotation of the southern end was towards the pole.

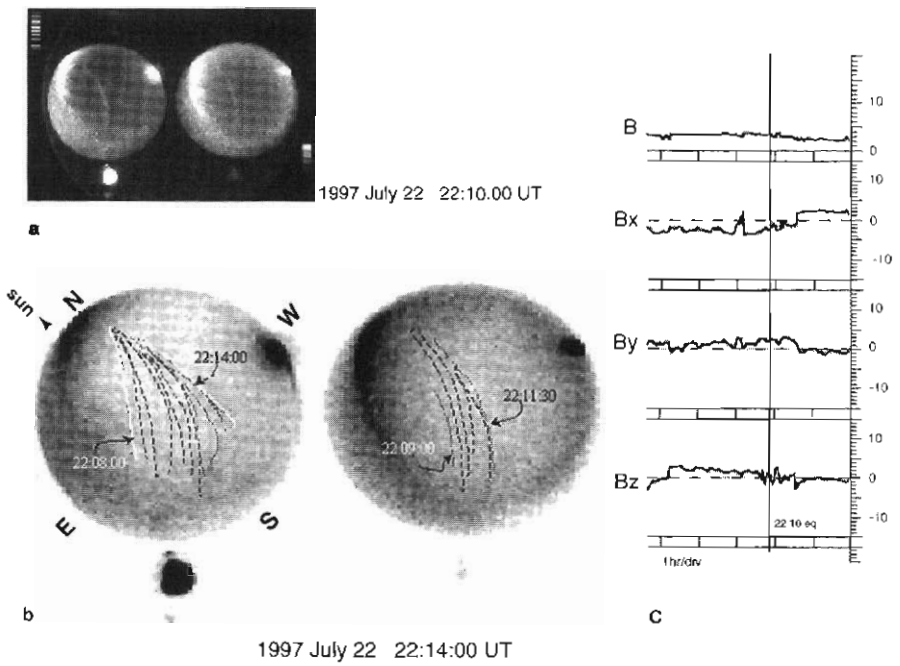


Figure 6. *22 July 1997. (a) raw data image; (b) arc positions; (c) WIND MFI data.*

16.3.4 7 July 1998

This arc, shown in Figure 7 moved north at about 300 ±40 m s^{-1} from 06:50 UT to arrive overhead at about 07:02 UT where it remained until fading into cloud/interference. This corresponds to 19:00 local time,

around dusk. The thumbnail Figure 7(a) shows the raw data at 07:05 UT. There is no information in the blue image due to the presence of the Moon. (A small spot visible 180° from the Moon in the red image is a reflection of the Moon from the inside of the spherical dome). In the composite image, Figure 7(b), the positions at two minute intervals from 06:50 until 07:16 UT are shown. This arc is again essentially aligned with the Sunward direction, the Sun being WSW at the time. The solar wind data are shown in Figure 7(c) with the projected time of the arc onset marked. The B_{tot} field is greater than the previous examples at around 10 nT and is more variable. B_x was between 8 and 5 nT for the preceding period, and remained positive. B_y was strongly negative at -10 nT, and had been negative for some time. B_z shows a sudden increase from around 1 to 6 nT at the start of this event. However as the viewing conditions were not perfect it is not possible to know if we have detected the actual birth of this event, or if it became visible after it was established.

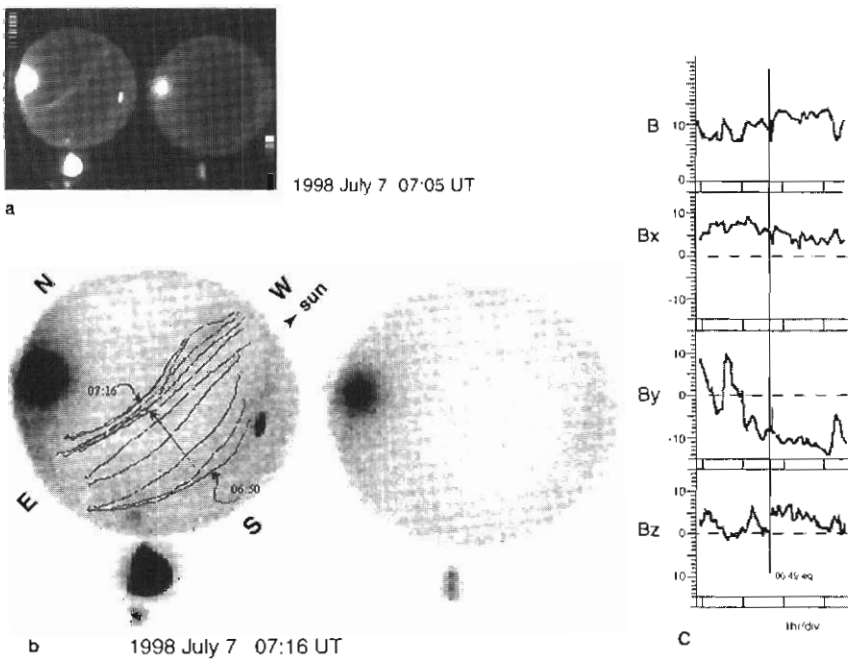


Figure 7. 7 July 1998. (a) raw data image; (b) arc positions; (c) WIND MFI data.

16.4 Discussion

The drift velocities of the first three arcs presented here were, 500 m s^{-1} poleward, 200 m s^{-1} anti-poleward, and a 'rotating' arc with up to 1000 m s^{-1} poleward motion, all recorded around local noon. The relation of the B_y direction to the poleward/anti-poleward drift motion in these specific examples is in agreement with the findings reported by Valledares *et al.* (1994), that the east-west component of the IMF

can control drift motion. The fourth event, from 07:00 UT (local dusk), shows a drift of 300 m s^{-1} northward which is unrelated to the poleward direction. This arc is aligned with the Sun direction. As stated before, all the polar cap arc examples presented here, and observed so far in the study are associated with northward IMF. This is expected from reports in the literature. The rotation of polar cap arcs is not well documented.

The four examples here were selected to show that there are several differing drift motion regimes. Detailed examination of the other identified events in our data is needed to establish trends. The images selected here represent some aspects of polar cap arc phenomena. Our data does contain aurora from the oval, and arcs that can be sufficiently long in extent to span the entire field-of-view. The straight or nearly straight localised arcs, aligned Sunward, or to the cusp (Valledares *et al.*, 1994) direction described here are unique to polar cap conditions. Specific examples that occur when reliable satellite data of IMF and solar wind are also available, are therefore potentially useful to examine in detail.

It is not clear whether these arcs occur on magnetic field lines that are 'open' to the interplanetary magnetic field, or field lines that map back into the magnetotail. Where the field lines map depends on the model of convection that applies, and more evidence is needed to justify either or both propositions. This paper presents preliminary observation results and the relation of these events with IMF B_y, B_x, and solar wind conditions will be expanded. Specific events will be related to data from magnetometers, riometers, and other photometers from the polar cap PENGUIn and AGO program, and ionospheric convection patterns.

16.5 Acknowledgments

The assistance of Antarctica New Zealand personnel at both Scott Base and Christchurch, New Zealand is gratefully acknowledged. This research was supported by a grant from the Australian Research Council, the University of Newcastle, and the Cooperative Research Centre for Satellite Systems with financial support from the Commonwealth of Australia through the CRC program.

References

Carlson Jr, H.C. and Egeland, A. (1995). The aurora and the auroral ionosphere. In: Kivelson, M.G. and Russell, C.T. (Eds.). *Introduction to Space Physics*. Cambridge University Press, New York.

Craven, J.D., Murphree, J.S., Frank, L.A. and Cogger, L. (1991). Simultaneous optical observations of transpolar arcs in the two polar caps. *Geophysical Research Letters*, 18: 2297–2300.

Feldstein, Y.I. and Starkov, G.V. (1967). Dynamics of the auroral belt and polar geomagnetic disturbances. *Planetary and Space Sciences*, 15. p. 209.

Frank, L.A. (1988). Dynamics of the near-Earth magnetotail: recent observations. In: Moore, T.E. and Wait, J.H. (Eds.). Modelling Magnetosphere Plasma. *Geophysical Monograph*, 44. p. 261.

Gussenhoven, M.S. and Mullen, E.G. (1989). Simultaneous relativistic electron and aurora particle access to the polar caps during interplanetary magnetic field B_z northwards: a scenario for an open field line source of auroral particles. *Journal of Geophysical Research*, 94: 17121–17132.

Hardy, D.A., Burke, W.J. and Gussenhoven, M.S. (1982). DMSP optical and electron measurements in the vicinity of polar cap arcs. *Journal of Geophysical Research*, 87: 2413–2430.

Sandholt, P.E., Farrugia, C.J., Moen, J. and Cowley, S.W.H. (1998). Dayside auroral configurations: responses to southward and northward rotations of the interplanetary magnetic field. *Journal of Geophysical Research*, 103: 20279–20295.

Peterson, W.K. and Shelley, E.G. (1984). Origin of the plasma in a cross-polar cap auroral feature (theta aurora). *Journal of Geophysical Research*, 89: 6729–6736.

Valledares, C.E., Carlson Jr, H.C. and Fukui, K. (1994). Interplanetary magnetic field dependency of stable Sun-aligned polar cap arcs. *Journal of Geophysical Research*, 99: 6247–6272.

Zhu, L., Schunk, R.W. and Sojka, J.J. (1997). Polar cap arcs: a review. *Journal of Atmospheric and Solar-Terrestrial Physics*, 59: 1087–1126.

17. MONITORING AURORALLY-GENERATED LARGE SCALE TRAVELLING IONOSPHERIC DISTURBANCES (TID'S) OVER A SUNSPOT CYCLE (1980–85)

L.A. Hajkowicz and B. Zimkowski

Department of Physics
University of Queensland
Queensland 4072 Australia
(email: hajkowicz@physics.uq.edu.au)

Abstract

The onset of auroral substorms is invariably associated with the generation of large-scale ionospheric disturbances (LSTID's) which modify the virtual height (h'F) of the F-region, primarily at sub-auroral locations. Five sub-auroral ionosonde stations, located in three largely different diurnal sectors: Australian-Asian (Hobart-Yakutsk), Indian Ocean (Kerguelen) and European (Uppsala-Dourbes) were used to monitor the aurorally-generated LSTID's over a sunspot cycle (1980–85). The virtual height enhancements (Δh'F) were analysed in conjunction with auroral substorm onsets deduced from the auroral electrojet index AE. Significant statistical results were obtained from the analysis of auroral-ionospheric data obtained during 473 disturbed days (for AE-index 400 nT). It is evident that substorm onsets are most frequent in a time interval: 1100–1900 UT, peaking at 1500 UT. Since LSTID's are strongly attenuated during daytime, due to the ion drag, the most advantageous geographical location for monitoring the disturbances is in the Australian-Asian sector which is then well into the night-time sector of the Earth, with the northern station Yakutsk being closest to the auroral source region. The other westward-displaced sectors are progressively less suitable for monitoring of LSTID's, as the propagation of LSTID's occurs in the sunlit sector of the Earth. Symmetrical propagation of LSTID's in both hemispheres is observed mainly in the equinoctial months.

17.1 Introduction

Large-scale ionospheric disturbances (LSTID's), which are an ionospheric manifestation of aurorally-generated atmospheric gravity-waves, have attracted considerable attention as a principal energy sink for auroral substorms (Walker *et al.*, 1988; Walker and Wong, 1993; Huang and Cheng, 1993; Prolss, 1993; Yeh *et al.*, 1992). Unlike other manifestations of auroral disturbances this wave phenomenon has a truly global character, spanning both hemispheres from high to equatorial latitudes and over a large range of longitudes. An extensive study has been undertaken on the global occurrence of LSTID's simultaneously launched from the northern and southern auroral ovals at the onset of auroral substorms. A particularly useful and readily available ionospheric tracer

of LSTID's was found to be the minimum height or the virtual height (h'F) of the F-region (Walker *et al.*, 1988). A sequential change in the ionospheric height is indicative of equatorward propagating trains of LSTID's in both hemispheres, with a typical propagation velocity between 500–1300 m s^{-1}, originating at the auroral locations characterised by a shell with L-value of about five (Hajkowicz, 1983a, 1983b; Hajkowicz, 1990; Hajkowicz, 1991a, 1991b; Hajkowicz and Hunsucker, 1987). The disturbances reach the highest amplitude at sub-auroral locations (Hajkowicz, 1991a, 1991b), rapidly decrease at mid-latitudes, and remain fairly constant at low-latitudes with a slight tendency to increase at the equator (Hajkowicz, 1991b).

A consistent factor in the propagation of LSTID's is their rapid attenuation in daytime which sometimes results in a gross asymmetry in the occurrences of disturbances in northern and southern hemispheres, following conjugate onsets of auroral substorms (Hajkowicz, 1990, 1991a). This attenuation is discussed in detail by Hajkowicz (1990) who attributed it to the ion drag retarding the gravity-wave propagation, associated with the high ionisation density of the F-region in daytime. The ion drag appears also to be responsible for pronounced differences in the occurrence of LSTID's at widely separated longitudinal (or diurnal) sectors, associated with the so-called 'universal time effect' of the onset of auroral substorms (Hajkowicz, 1992). The effect was only studied at a relatively high sunspot number (1980–82), but because of its global significance was also extended to the low sunspot activity interval (1983–85) in the present considerations. The present study, which completes the previous results, clearly demonstrates the preference of certain geographical locations in monitoring the global morphology of large scale gravity-waves.

17.2 Results

The ionospheric height enhancements Δh'F were obtained by subtracting monthly median values of h'F from the current height variations, thus removing the diurnal thermospheric wind effect. Numerous examples of sequential ionospheric height enhancements Δh'F, occurring progressively later as LSTID's propagated towards the equator, has been presented in previous reports (e.g. Hajkowicz, 1992). In all these cases the virtual height enhancement invariably followed an onset of auroral substorms as inferred from a rapid rise in the auroral electrojet index, AE. It is evident that a statistical study of h'F enhancement, particularly at locations close to the source region (Hajkowicz, 1991b), is of considerable importance in the long-term evaluation of the ionospheric response to substorm activity. In the initial phase of the project six sub-auroral ionosonde stations, as specified in Table 1, were used in the analysis. It became evident that average auroral activity had a distinct diurnal distribution, peaking between 1300–1900 UT and the most distinct height rises tended to occur for the stations which were then at night-time, i.e. in the Australian-Asian sector (Hobart-Yakutsk). Other stations, west of this sector and towards the sunlit region of the Earth (i.e. Kerguelen, Uppsala, Dourbes and Ottawa) showed progressively

smaller LSTID's activity. Thus, the average height enhancements were smaller in the European sector (Dourbes-Uppsala) and were very much reduced in the North American sector (Ottawa) which was in full daylight during the period of maximum auroral activity.

***Table 1**. Ionosonde stations used in the height analysis. Average night-time (at an altitude of 300 km), geographic latitude and East longitude, invariant latitude and L-value are indicated.*

	Geographic		Invariant		Night-time interval (UT)			
Station	Lat. (deg.)	Long.	Lat. (deg.)	L	Summer	Autumn	Winter	Spring
Hobart	-42.9	147.3	53.6	2.8	12-17	10-19	08-20	10-19
Yakutsk	62.0	129.6	55.1	3.1	08-22	12-18	Day	12-18
Kerguelen	-48.8	70.0	58.2	3.6	18-20	14-23	13-01	14-23
Uppsala	59.8	17.6	56.6	3.3	16-05	20-02	Day	20-02
Dourbes	50.8	4.4	48.5	2.3	18-06	20-04	23-01	20-04
Ottawa	45.4	284.3	57.7	3.5	01-13	01-09	03-07	01-09

The same stations (except for Ottawa for which much of the data was missing) were used in the analysis of ionospheric response in the sunspot minimum period (1983–85). Examples of ionospheric responses to individual substorms over a wide range of longitudes have been presented in Figuress 1 and 2 by Hajkowicz (1992). The examples clearly indicate a delay effect between the occurrence of auroral disturbances and the subsequent ionospheric height enhancements, characteristic of LSTID's propagation. A large number of individual cases was analysed, resulting in an average variation of AE and the associated height enhancement $\Delta h'F$ for specific seasons. The high degree of correlation of the ionospheric responses (in the Asian-Australian sector) with the auroral electrojet surges can be seen in the diurnal variations of $\Delta h'F$ (derived from averaging responses at Hobart and Yakutsk, symbolised as $\Delta H'F$) for selected seasons in Figure 1. A number (indicated in brackets) of disturbed days were used to derive the average $\Delta h'F$ and corresponding AE. It can be also seen that in all cases $\Delta h'F$ for Hobart (H) and Yakutsk (Y) closely resembles the average trend in the auroral electrojet activity peaking between 1200–1700 UT, within the local night-time for these stations. The responses at Yakutsk and Hobart were of similar duration in the equinoctial months, with similar night-time (at the height of the F-region) periods at both stations. In summer (northern winter) the response was more extended at Yakutsk than at Hobart as the former station was in the winter hemisphere (Figures 1(b) and 1(d)), characterised by long night-time, facilitating propagation of LSTID's in

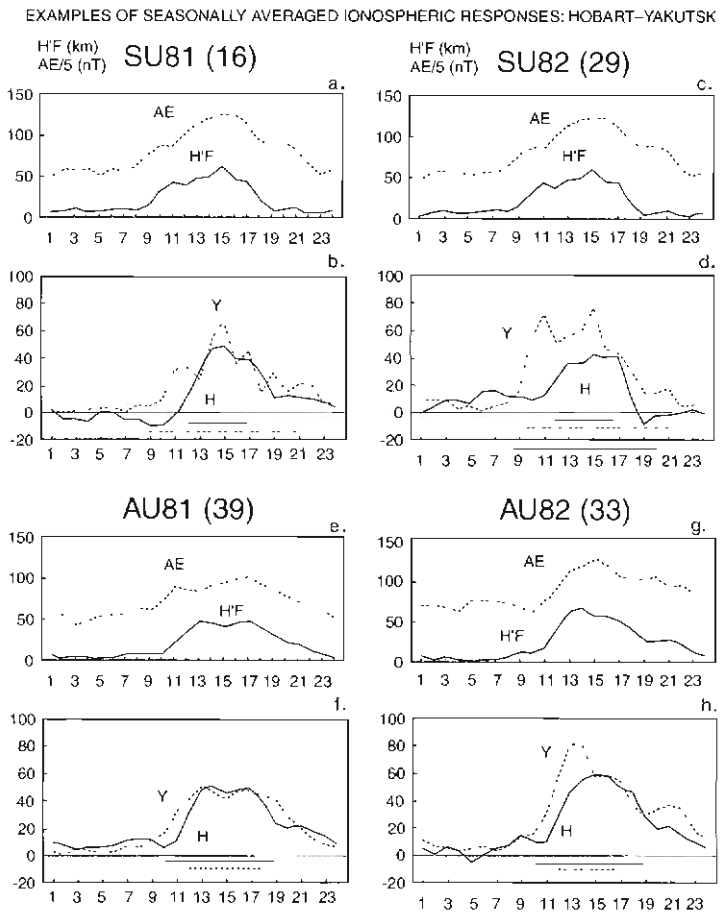


Figure 1. *Examples of seasonally averaged ionospheric responses for individual stations: Hobart (H-solid line) and Yakutsk (Y-broken line) and the corresponding average responses for the Australian-Asian sector ($\Delta H'F$ – solid line) and the associated AE-index (scaled down by a factor of five). The southern seasons and years (e.g. SU81, AU81 refers to summer 1981 and autumn 1981) are indicated (the number of disturbed days used for averaging are indicated in the brackets). The solid and broken lines indicate the corresponding night-time (at an altitude of 300 km) for each season.*

the northern hemisphere. It is also evident that the responses at Yakutsk tend to be of a larger amplitude than those at Hobart as the former station is closer to the source region (larger L in Table 1).

Note that the average ionospheric responses ($\Delta H'F$) in the Australian-Asian sector follow almost exactly the trend in the auroral electrojet index in autumns 1981 and 1982 (Figures 1(e) and 1(g)). This has been noted for individual disturbances by Hajkowicz (1991b) who reported that the correlation coefficient between AE and the corresponding height enhancements was on average 0.7, on occasions reaching 0.9.

The complete average ionospheric response for all seasons (Autumn and Spring are combined in the equinoctial period) can be seen in Figure 2 for a large number of disturbed days in each season at sunspot maximum (1980–82) and minimum (1983–85). Altogether, 473 disturbed days, with a minimum onset value of AE = 400 nT, were used in the analysis.

It is now evident that the universal time effect in the diurnal preference in the occurrence of substorm onsets is maintained throughout the solar cycle for all seasons, except in winter (Figures 2(e) and 2(g)) where the substorm onsets are more evenly distributed throughout the day. Thus the majority of auroral surges, as signified by the rapid variations in AE, were in an interval 1100–1900 UT coinciding with night-time at Hobart (local midnight 1400 UT) and at Yakutsk (local midnight at 1500 UT). It should be noted that during southern summer (northern winter), particularly during sunspot maximum, the response at Yakutsk is wider and larger than at Hobart as Yakutsk has then a long night-time, resulting in more favourable propagation conditions for LSTID's (also demonstrated for the selected summers in Figure 1). This was particularly evident at sunspot maximum (Figures 2(a) and 2(b)). In southern winter (northern summer) Yakutsk is in continuous daytime and shows less extended ionospheric response than Hobart (cf. Figures 2(e) and 2(f), and 2(g) and 2(h)). It is of considerable interest that LSTID's recorded in this season at Yakutsk propagate through the sunlit atmosphere, with most of the disturbances reaching the station close to local midnight (Figures 2(f) and 2(h)), i.e. at the time when the Sun reaches the lowest elevation at this station. Generally the response at Yakutsk tends to be larger in magnitude than that at Hobart, particularly at equinox at sunspot minimum, and in the post-midnight sector (Figure 2(m) and 2(n)). It is clearly evident (from the large statistical samples) that symmetrical propagation of LSTID's in both hemispheres, following of the onset of auroral substorms, is characteristic of the equinoctial periods for the solar cycle (Figure 2(k) and 2(l), and 2(m) and 2(n)). It should also be noted that a sharp peak in the ionospheric response at Yakutsk is observed at almost the same time, 1500–1600 UT (i.e near local midnight), in northern winter (Figures 2(b) and 2(d)) and in the equinox (Figure 2(l) and 2(n)) for sunspot maximum and minimum. This sharp peak is mostly missing in the responses recorded at Hobart.

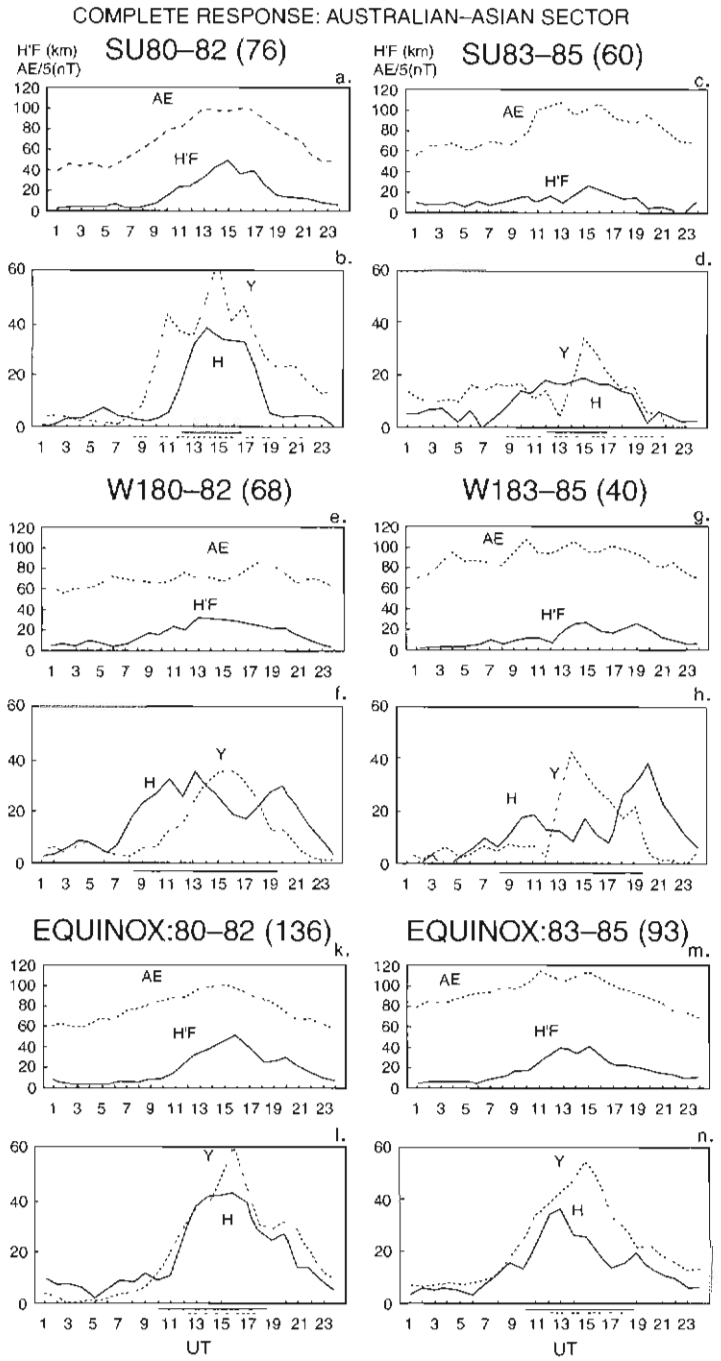


Figure 2. *The complete ionospheric responses in the Australian-Asian sector, averaged for each season at sunspot maximum (1980–82) and minimum (1983–85). (The symbols are as defined in Figure 1).*

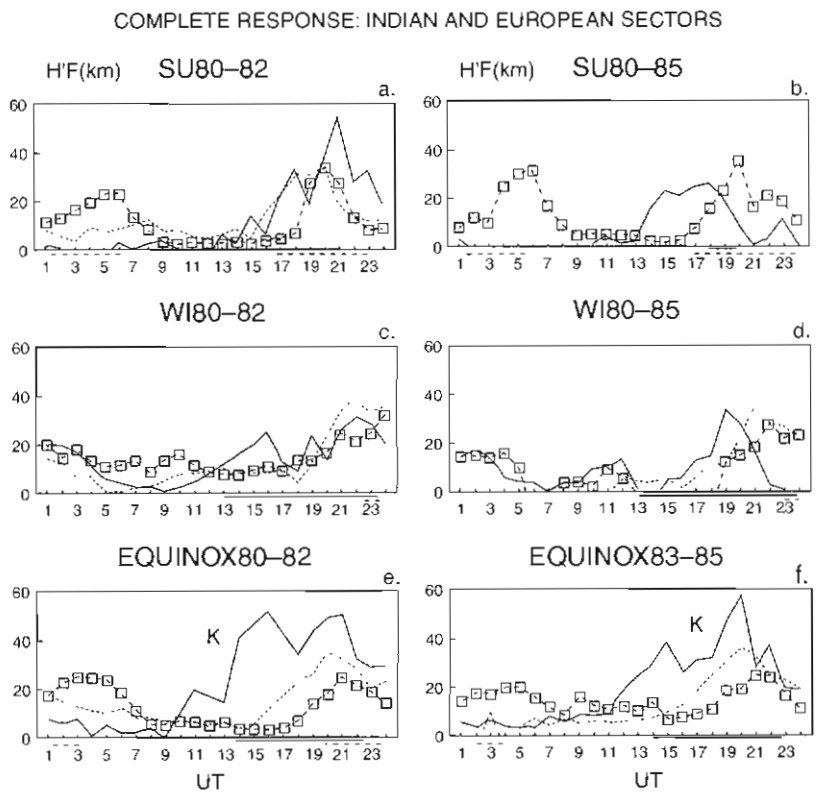


Figure 3. *The complete ionospheric responses in the Indian (for Kerguelen-K) and European (for Uppsala-U and Dourbes-D) sectors for different seasons at sunspot maximum and minimum. (The line symbols for three stations are given in e and f).*

The ionospheric response is less distinct in two other longitudinal sectors: Indian Ocean (Kerguelen– K) and Europe (Uppsala–U and Dourbes–D) positioned towards the evening side of the auroral electrojet maximum (Figure 3). The maximum frequency of auroral substorm onsets occurs close to sunset at Kerguelen but the peak of auroral activity is at local afternoon at Uppsala and Dourbes. Consequently, a large proportion of aurorally-generated LSTID's are attenuated in these sectors, resulting in a narrower and weaker ionospheric response than that for the Australian-Asian sector. It should be noted that the response at Kerguelen is quite high at equinox in the post-sunset sector (particularly during sunspot

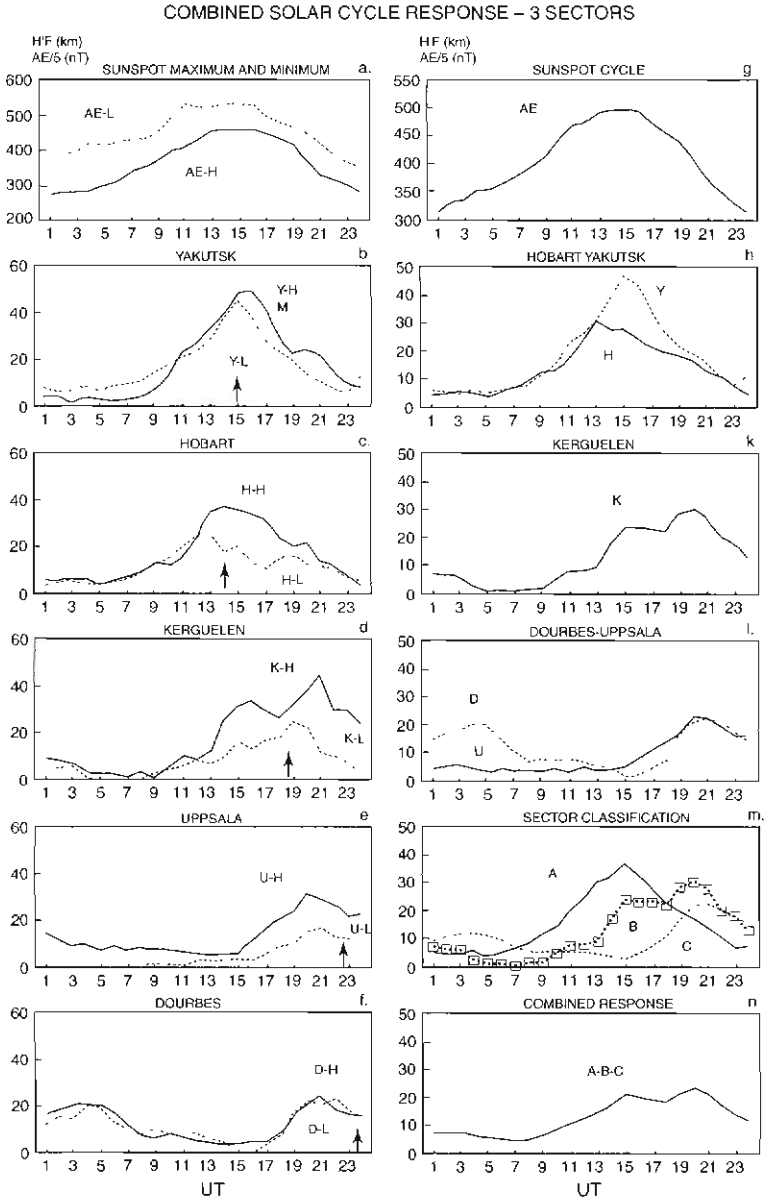


Figure 4. *The sunspot maximum and minimum (a–f) and combined sunspot cycle (g–n) responses. The symbols AE–L, AE–H, Y–H etc. refer to the average values of AE-index at low and high sunspot numbers, average Δh'F for Yakutsk for high sunspot numbers. The arrows indicate the local midnight.*

maximum, Figure 3(e)) as this station, positioned at L = 3.6, is nearest (from all the stations) to the auroral source region. The European sector shows a much weaker response both in magnitude and duration over the sunspot cycle. In the equinoctial months there is a consistent decline in the ionospheric response, from Kerguelen to Dourbes at sunspot maximum and minimum (Figures 3(e) and 3(f)). This again is in agreement with the attenuation of LSTID's as they propagate from the source region to lower latitudes. Finally, in summer most LSTID's propagation at higher latitudes (particularly at Kerguelen, Figure 3(a)) takes place in the sunlit atmosphere as was the case for Yakutsk. The comparative average ionospheric response over the sunspot maximum and minimum, and over the entire cycle are given in Figure 4. It is evident that the trend in the onset of auroral substorms shows the characteristic universal time effect, with substorm activity mainly concentrated between 1100–1900 UT (with a peak at about 1500 UT), over the solar cycle (Figures 4(a) and 4(g)). The peak in AE is closely followed by a peak in $\Delta h'f$ at Yakutsk (Figure 4(b)), at sunspot minimum (Y–L) and maximum (Y–H). A sharp maximum is reached at Yakutsk at local midnight for the entire sunspot cycle (Figure 4(h)). On the whole the response at Yakutsk is more pronounced than at Hobart (Figure 4(h)), particularly in the post-midnight sector. The response is pronounced at Kerguelen being the same magnitude as Hobart but shifted away from the center of auroral activity, towards local night-time (Figures 4(d) and 4(k)). The European sector is characterised by a much weaker response at Uppsala and Dourbes (Figures 4(e) and 4(f)) as compared to other sectors (Figure 4(m)).

It is evident that the response at sunspot minimum is consistently lower than that at sunspot maximum for all the sub-auroral stations (L = 2.8; Figures 4(b)–4(e)) despite the average associated AE-index is larger at sunspot minimum than maximum (Figure 4(a)).

Figure 4(m) is of particular interest as it gives a concise classification of the three sectors according to their ability to monitor aurorally-generated LSTID's. Thus, the most effective sector in detecting the disturbance is the Australian-Asian region (A), followed by the Indian Ocean region (B) and then the European region (C). The combined response (averaged A, B and C) provides monitoring facilities of LSTID's from about 1200 to 2300 UT (Figure 4(n)).

17.3 Results and discussion

The statistical analysis of ionospheric response (using 473 disturbed days), associated with the propagation of large scale ionospheric disturbances (LSTID's), was conducted at three largely different longitudinal sectors at northern and southern sub-auroral locations. For the first time it was possible to establish the longitudinal preference in monitoring LSTID's over a sunspot cycle. The following is evident:

(i) The substorm onsets, as inferred from the auroral electrojet surges (AE-index 400 nT), are most frequently encountered in a time interval

1100–1900 UT, peaking at about 1500 UT. This trend in auroral activity appears to be present in other solar cycles as discussed by Hajkowicz (1992).

(ii) The propagation of LSTID's is affected by ion drag and critically depends on the solar illumination conditions. Thus, the Australian-Asian sector which is close to local midnight during the peak of auroral activity, records LSTID's considerably more frequently than other westward-displaced sectors. The results obtained for the American sector at sunspot maximum (Hajkowicz, 1992) give a global classification of geographic locations suitable for monitoring LSTID's (from the most to the least suitable): Australian-Asian sector (Hobart, Yakutsk), Indian Ocean sector (Kerguelen), European sector (Uppsala, Dourbes) and American sector (Ottawa).

(iii) Since monitoring LSTID's is also most effective at sub-auroral locations, the Asian station Yakutsk (L = 3.1), appears to be the most suitable location for monitoring LSTID's. (The locations shifted even closer to the source region would be impractical due to frequent absorption of ionosonde signals). The southern equivalent of this station in Australia, Hobart, also has an important role in monitoring simultaneous LSTID's propagating equatorwards from the southern auroral oval.

(iv) There is a consistent decline in the amplitude of LSTID's from sunspot minimum to maximum for all sub-auroral stations (L = 2.8), especially in the post-midnight sector. This decline was particularly noticeable for the southern sub-auroral station Hobart.

The results presented support previous findings on attenuation of large scale gravity waves due to ion drag, which retard the propagation of the disturbance in daytime. It should be noted however that this attenuation does not prevent the gravity wave from reaching a sub-auroral station, such as Yakutsk, even at daytime as can be seen from the relatively large response at Yakutsk during northern summer when both the station and source region are in continuous daylight (Figure 2(f)). This is supported by a case study (during the great storm of 13 March 1989, by Hajkowicz, 1991a) when a large auroral surge was followed by a global height enhancement, including a large height rise at Uppsala which was in full daylight. The attenuation along the disturbance propagation path, is however, very rapid during daytime. Little or no disturbance is observed at lower latitude stations (Hajkowicz, 1990a, 1991a). On the other hand, LSTID's propagate almost with a constant amplitude to low-latitudes at night-time (Hajkowicz and Hunsucker, 1987).

A separate and unknown aspect of the present study concerns the consistent decline in the amplitude of LSTID's during sunspot minimum. This can be associated with a decline in the median height h'F from sunspot maximum to minimum. For example, a study of the median height of the F-region in the Asian–Australian sector over a sunspot cycle indicates a gradual decrease in the background h'F at all seasons

(Hajkowicz, 1991c). Such background height variations may have a substantial effect on the amplitude of LSTID's as the gravity-wave amplitude grows exponentially with altitude, i.e. with the height of the F-region. This could lead to lower amplitudes of LSTID's at sunspot minimum as observed.

17.4 Acknowledgements

The ionosonde data were supplied by the IPS Radio and Space Services, Sydney, Australia; The World Data Center B2, Moscow, Russia; The Institut Royal Meteorologique de Belgique, Brussel, Belgium; The Institute of National Defence, Stockholm, Sweden and The Centre National d'Etudes des Telecommunications, Lannion, France. We are grateful to Mrs D.J. Dearden for her assistance with the data analysis and to Dr G.G. Bowman for his discussion on some aspects of this work.

References

Hajkowicz, L.A. (1983a). Auroral riometer absorption and the F-region disturbances observed over a wide range of latitudes. *Journal of Atmospheric and Terrestrial Physics*, 45:175–179.

Hajkowicz, L.A. (1983b). Conjugate effects in the generation of travelling ionospheric disturbances (TIDs) in the F-region. *Planetary and Space Science*, 31: 1409–1413.

Hajkowicz, L.A. (1990). A global study of large scale travelling ionospheric disturbances (TIDs) following a step-like onset of auroral substorms in both hemispheres. *Planetary and Space Science*, 38: 913–923.

Hajkowicz, L.A. (1991a). Global onset and propagation of large scale travelling ionospheric disturbances as a result of the great storm of 13 March, 1989. *Planetary and Space Science*, 39: 583–593.

Hajkowicz, L.A. (1991b). Auroral electrojet effect on the global occurrence pattern of large scale travelling ionospheric disturbances. *Planetary and Space Science*, 39: 1189–1196.

Hajkowicz, L.A. (1991c). Median ionospheric height variations over a sunspot cycle in the Australian-Japanese longitudinal sector. *Planetary and Space Science*, 39: 1607–1615.

Hajkowicz, L.A. (1992). Universal time effect in the occurrences of large-scale ionospheric disturbances. *Planetary and Space Science*, 40: 1093–1099.

Hajkowicz, L.A. (1995). Ionospheric response to auroral substorms during sunspot maximum (1980–82). *Annales Geophysicae*, 13: 95–104.

Hajkowicz, L.A. and Hunsucker, R.D. (1987). A simultaneous observation of large-scale periodic TIDs in both hemispheres following an onset of auroral disturbances. *Planetary and Space Science*, 35: 785–791.

Huang, Y.-N. and Cheng, K. (1993). Ionospheric disturbances around East Asian region during the 20 October 1989 magnetic storm. *Journal of Atmospheric and Terrestrial Physics*, 55: 1009–1020.

Prolss, G.W. (1993). On explaining the local time variation of ionospheric storm effects. *Annales Geophysicae*, 11: 1–9.

Walker, G.O. and Wong, Y.W. (1993). Ionospheric effects observed throughout East Asia of the large magnetic storm of 13-15 March 1988. *Journal of Atmospheric and Terrestrial Physics*, 55: 1009–1020.

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

Yeh, K.C., Liu, K.H. and Conkright, R.O. (1992). The global behaviour of the March 1989 ionospheric storm. *Canadian Journal of Physics*, 70: 532–543.

18. IONOSPHERIC SLAB THICKNESS AND TOTAL ELECTRON CONTENT DETERMINED IN AUSTRALIA

A.M. Breed[1] and G.L. Goodwin[2]

[1] Atmospheric and Space Physics
Australian Antarctic Division
Kingston Tasmania 7050 Australia
(email: anthony.breed@aad.gov.au)

[2] School of Physics and Electronic Systems Engineering
The Levels Campus
University of South Australia,
Mawson Lakes SA 5095 Australia

Abstract

Ionospheric Total Electron Content (TEC) determined from July 1991 to June 1995 is examined from southern Australia using Global Positioning System (GPS) satellite data recorded by the Defence Science and Technology Organisation at Salisbury, South Australia. Midday TEC was observed to increase with sunspot number. Protonospheric TEC was estimated from the difference between GPS TEC and TEC determined from Faraday rotation measurements. Ionospheric slab thickness from 1991 to 1993 (when protonospheric TEC was negligible) showed no substantial latitudinal variation for each season and year. Midday slab thickness is greater in summer and is greater nearer to solar maximum.

18.1 Introduction

TEC is the total electron content (electrons m^{-2}) of a vertical column of 1 m^2 cross-section from the height of a GPS satellite (20 000 km) to ground level. TEC is measured in TEC Units (TECU) where 1 TECU = 10^{16} electrons m^{-2}. The present set of measurements of TEC in South Australia at Salisbury (latitude 34.77°S, longitude 138.63°E) commenced in July 1991 to satisfy a need for extended TEC data in the southern hemisphere. This was desirable as a reliable basis for providing the correction for the ionospheric time difference in the reception of radio transmissions at 1227.60 MHz compared with 1575.42 MHz from Global Positioning System (GPS) satellites (1 ns of time difference corresponds to 2.8 TECU and 0.3 metres of signal path from satellite to receiver). Ionospheric corrections of approximately 10 to 50 TECU for (vertical) TEC are inferred from Figure 2 which correspond to approximately one to five metres in the vertical component of signal path. Corrections of this magnitude are usual in the position determination of a GPS receiver at ground level.

Errors in TEC measurements include those due to receiver offset bias and satellite offset bias which are time delay errors produced in the hardware

of the receiver and satellite respectively between signals at the two GPS frequencies. These errors were removed as much as practicable from the TEC data. A Rockwell UH60 GPS receiver was used until April 1994, and subsequently an Ashtech 12 receiver (which required a large offset bias correction).

18.2 Corrected GPS TEC

A recent (midnight, autumn 1997) TEC versus latitude plot is shown in Figure 1. The computed TEC values (representing the ionosphere and most of the protonosphere) are too large by approximately 23 TECU due mainly to a big offset bias error of the Ashtech 12 receiver. When a correction derived from geometry is applied to the plot in Figure 1, it becomes a straight line (denoted 'corrected GPS TEC' in the figure); i.e. the correction removes the pronounced maximum around latitude 35°S, the receiver latitude.

Goodwin and Breed (1999) provided the correction as follows. TEC is equal to slant TEC × cos(i) where slant TEC is measured along the signal path from satellite to receiver, and (i) is the incidence angle of the ray on the ionosphere assumed to be at a median height of 400 km. Observations are made at latitude 35°S. It is assumed that satellites with subionospheric points at latitudes of 45°S and 25°S have the same 'slant correction factor', $\cos(i_1)$. Satellites at latitude 35°S were shown by Goodwin and Breed (1999) to have an average 'slant correction factor', $\cos(i_2)$ which was greater than $\cos(i_1)$. (In fact, i_2 corresponds to a set of subionospheric points at latitude 35°S which have a mean difference in longitude from that of the observation point equivalent to a difference of approximately 3° of latitude). The above considerations provide the ratio of the TEC correction at 45°S and 25°S to that at 35°S.

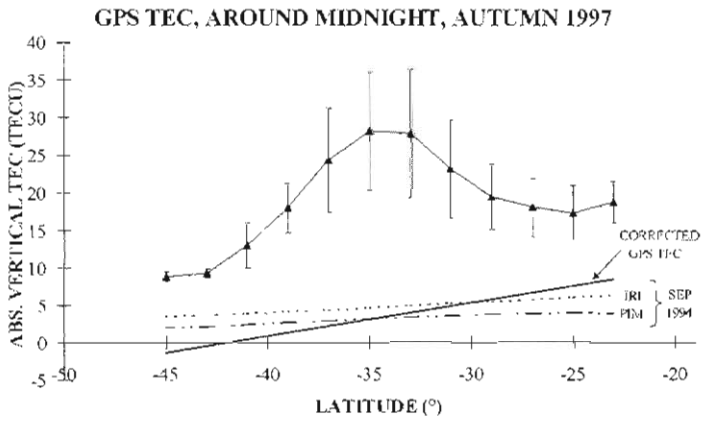


Figure 1. *A comparison of GPS TEC versus latitude measured at Salisbury with model prediction from IRI and PIM.*

Furthermore, it is assumed that the slope of the corrected GPS TEC versus latitude plot is the same as the slope for the original measurements at 45°S and 25°S. The above information is sufficient basis to provide a corrected TEC versus latitude plot shown in Figure 1.

Model values for only ionospheric TEC (September 1994) are shown; these values were computed for a time when (compared with the GPS TEC values) the model values would be approximately 3 TECU higher due to slightly greater solar activity, but would be 2 TECU less than for the corresponding GPS TEC values because of the absence of protonospheric TEC in the models. The corrected GPS TEC and model values should therefore be similar in magnitude, which is demonstrated in Figure 1. Two different TEC models are used: The International Reference Ionosphere (IRI 1990) (Bilitza, 1990) and the Parameterised Ionospheric Model (PIM, Version 1.4, February 1996) (Daniell *et al.*, 1995). The IRI allows TEC predictions to a maximum height of 3000 km while PIM allows the calculation of electron density (and hence TEC) to an altitude of 1600 km.

In Figure 1 the plot for the corrected GPS TEC has a greater slope than that predicted from IRI and PIM. It is possible that appreciable observational errors, such as those around latitude 25°S, could have led to an error in the 'corrected slope' of a magnitude that could account for the discrepancy. The apparently negative corrected GPS TEC values at higher latitudes could also be attributable to appreciable observational errors.

18.3 Protonospheric TEC determined from GPS TEC

Measurements of GPS TEC are immediately useful in determining the retardation of a radio signal from a GPS satellite to a ground receiver, so that a correction may be applied in determining a receiver's location. However, in using GPS TEC for ionospheric purposes, it is normally desirable to estimate, and deduct, the protonospheric TEC component.

The protonosphere, or plasmasphere, is the region above the O^+/H^+ transition height. The protonosphere is normally considered to occur from 2000 km to 35 000 km, the height of the plasmapause, with only a small TEC contribution to the protonosphere occurring between 20 000 km (the height of GPS satellites) and 35 000 km. The protonosphere corresponds to approximately 10% of the combined ionosphere-protonosphere in the daytime when the electron density in the ionosphere is high, and 40–50% of the night-time TEC when the ionospheric electron density is low. These percentages are representative of the protonosphere over a solar cycle; the percentages will normally be lower near solar maximum and higher near solar minimum. The protonosphere has no internal ionisation production. It is enhanced by an upward diffusion of ionisation from the ionosphere during the day, and depleted by a downward diffusion at night.

The contribution of protonospheric TEC to the total GPS TEC is considered in terms of the diurnal ratio of the daytime maximum TEC to the night-time minimum TEC as follows. Using the diurnal TEC ratios in Table 1, a comparison is made between the TEC measurements of Essex (1978) determined using the Faraday rotation technique for the ionosphere (up to 2000 km) and the present GPS TEC measurements (up to 20 000 km). Both sets of measurements are for southern Australia in the declining phase of a solar cycle.

Table 1. *Essex (1978); Diurnal TEC ratios* obtained from Faraday rotation measurements of TEC recorded in Melbourne.*

TEC ratios	1971	1972	1973	1974	Average diurnal TEC ratios for ionosphere
Summer	4.1	4.3	5.7	6.0	5.0
Equinox	6.4	5.9	6.2	6.3	6.2
Winter	5.2	5.4	4.8	5.2	5.2
Mean	5.2	5.2	5.6	5.8	5.5

Breed and Goodwin (2000); Diurnal TEC ratios* obtained from GPS measurements of TEC recorded at Salisbury (near Adelaide).

TEC ratios	1991	1992	1993	1994	1995	Average diurnal GPS TEC ratios	Average diurnal TEC ratios for ionosphere**
Summer	3.4	6.0	1.8	3.6		3.7	6.7
Equinox	5.1	4.8	2.9	2.0		3.7	6.7
Winter	4.4		1.7	2.7	2.6	2.9	5.2
Mean	-	-	-	-	-	3.4	6.2

* The diurnal TEC ratio is the ratio of daytime maximum TEC to night-time minimum TEC.

** Diurnal TEC ratio for ionosphere = (90/50) x diurnal GPS TEC ratio.

The ionospheric TEC is expected to be 90% of the GPS TEC for the daytime maximum, and approximately 50% for the night-time minimum. In Table 1 the present average diurnal GPS TEC ratios are

therefore multiplied by 90/50 to produce the approximate average TEC ratios for the ionosphere. The reasonable measure of agreement between 6.2 (with an estimated error of approximately ±12%) for the present mean diurnal ratios and 5.5 for those of Essex (1978) indicate that ionospheric TEC can be determined by removing the protonospheric TEC from GPS TEC.

Figure 2 shows midday TEC values plotted seasonally against smoothed values of sunspot number from 1991 to 1995. The regression coefficient, R^2, is indicated for each seasonal plot. TEC is related to sunspot number, S, by an equation (Huang, 1978) of the form

$$TEC = A + B \times S$$

where A and B are constants dependent on season and time of day.

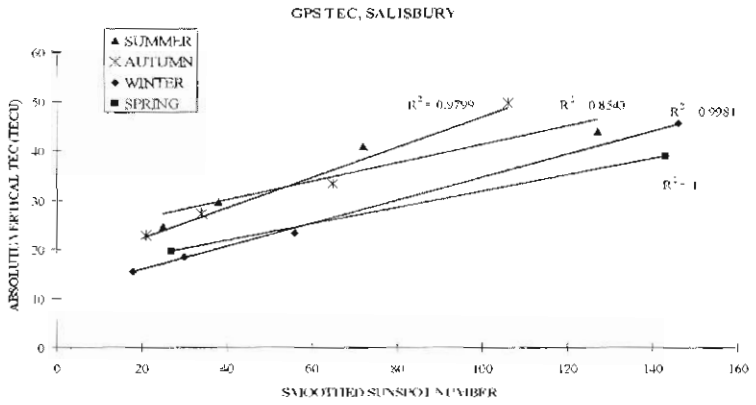


Figure 2. *Midday TEC values are plotted against the smoothed values of sunspot number for the specified season in each year. Regression coefficient, R2, is indicated for each seasonal plot.*

Midday values of A and B shown in Table 2 were derived for the present GPS TEC data, using a linear regression. These are compared with ionospheric TEC data of Huang (1978) obtained at a similar latitude from Faraday rotation observations up to a height of 2000 km. Since the parameter A is a TEC offset, it follows that the inclusion of protonospheric TEC in the present GPS TEC data has the effect of increasing this offset. It follows that the difference between the values of A for the two data sets is a rough estimate of the protonospheric TEC between 2000 km and 20 000 km heights. In Table 2 the difference between the present median A-value (for seasonal midday values in

1991–1995) and that of Huang (1978) is 15.5 - 8.5 = 7 TECU, which is a rough estimate of the protonospheric TEC at midday. Another example is the difference between the summer A-values which is 22.5 - 9.5 = 13 TECU for the summer midday protonosphere.

Table 2. *Approximate values of constants A and B for the equation, TEC = A + B x S are listed for the seasons; S is the sunspot number. (Errors associated with the seasonal Salisbury values are estimated at ±10%).*

	Salisbury (34.6°S) (Present work)		Sagamore Hill (38.6°N) Huang (1978)	
Seasons	A	B	A	B
Summer	22.5	0.19	7 – 12	0.1 – 0.15
Autumn	15.9	0.31	10 – 15	0.2 – 0.25
Winter	11.0	0.23	5 – 10	0.25 – 0.3
Spring	15.1	0.17	5 – 10	0.25 – 0.3
Median	15.5	-	8.5	-

Figure 3 shows the diurnal variation in protonospheric TEC at 2000 km to 20 000 km determined from the difference between GPS TEC and Faraday rotation TEC recorded near Salisbury in December (summer) 1992, in the declining phase of the solar cycle. The median protonospheric TEC is 7.2 TECU. Similarly, the summer midday protonosphere has a TEC of 15 TECU, which is in reasonable agreement with the rough estimate of 13 TECU determined from Table 2.

Earlier workers (Almeida *et al.*, 1970; Davies *et al.*, 1976; Kersley *et al.*, 1978) calculated protonospheric TEC from measurements of the retardation of transmissions along fixed paths to ground receivers from the geostationary satellite, ATS-6, located over the equator at an altitude of 35 000 km. They combined two techniques, as follows. In the Faraday rotation technique, slant TEC was measured up to 2000 km. In the differential phase technique, slant TEC was measured up to 35 000 km. It follows that the difference between these two measurements is the slant TEC of the protonosphere. They recorded slant protonospheric TEC in Aberystwyth (U.K.), Hamilton (eastern U.S.A.) and Boulder (middle North America).

In Figure 4, the diurnal variation of (vertical) protonospheric TEC observed at Salisbury (near Adelaide) is plotted. A line of best fit is sketched through the plotted points which each have an estimated accuracy of ±8 TECU. The 'dip' in TEC around 0700 UT could be attributed to the limitations in the accuracy of the data.

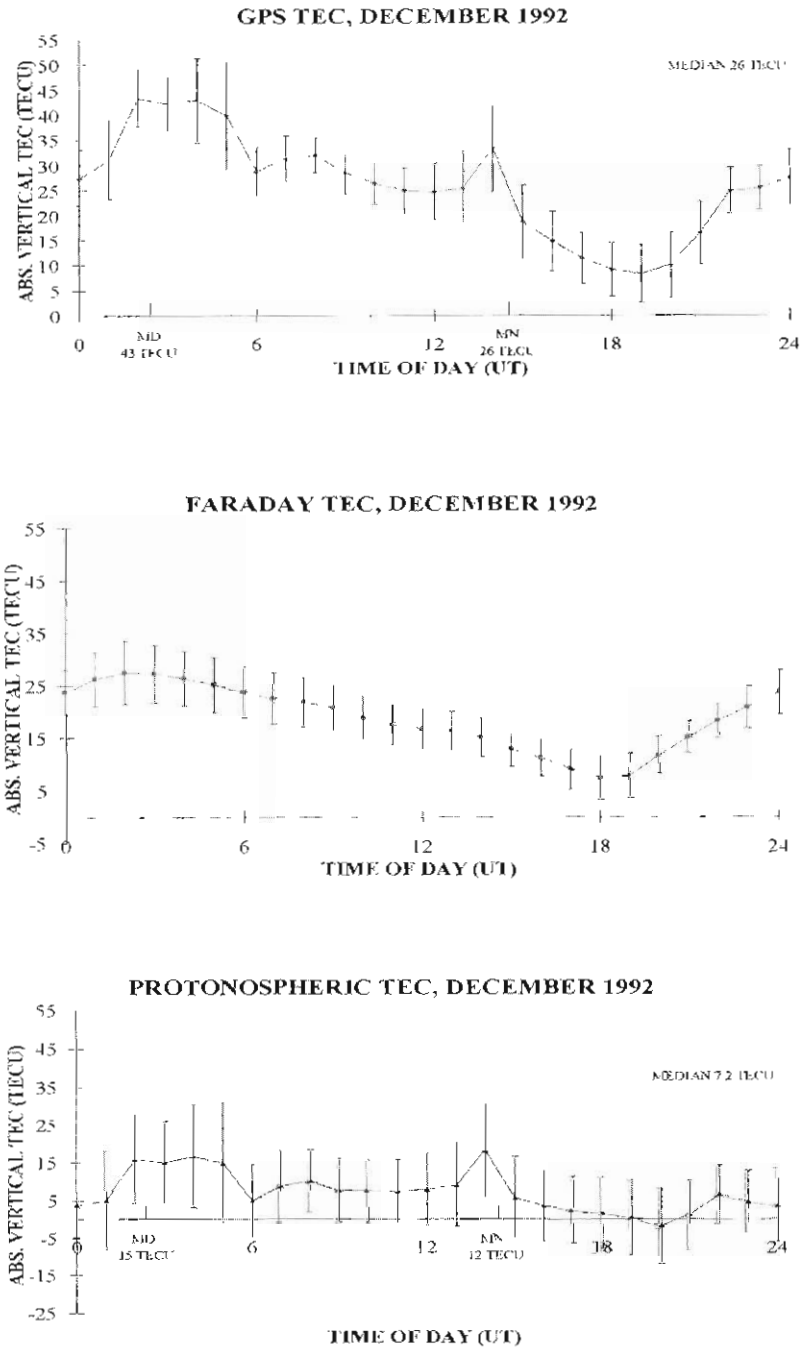


Figure 3. *Hourly average TEC values derived from GPS and Faraday rotation measurements at Salisbury and the calculated protonospheric electron content for December 1992 are shown.*

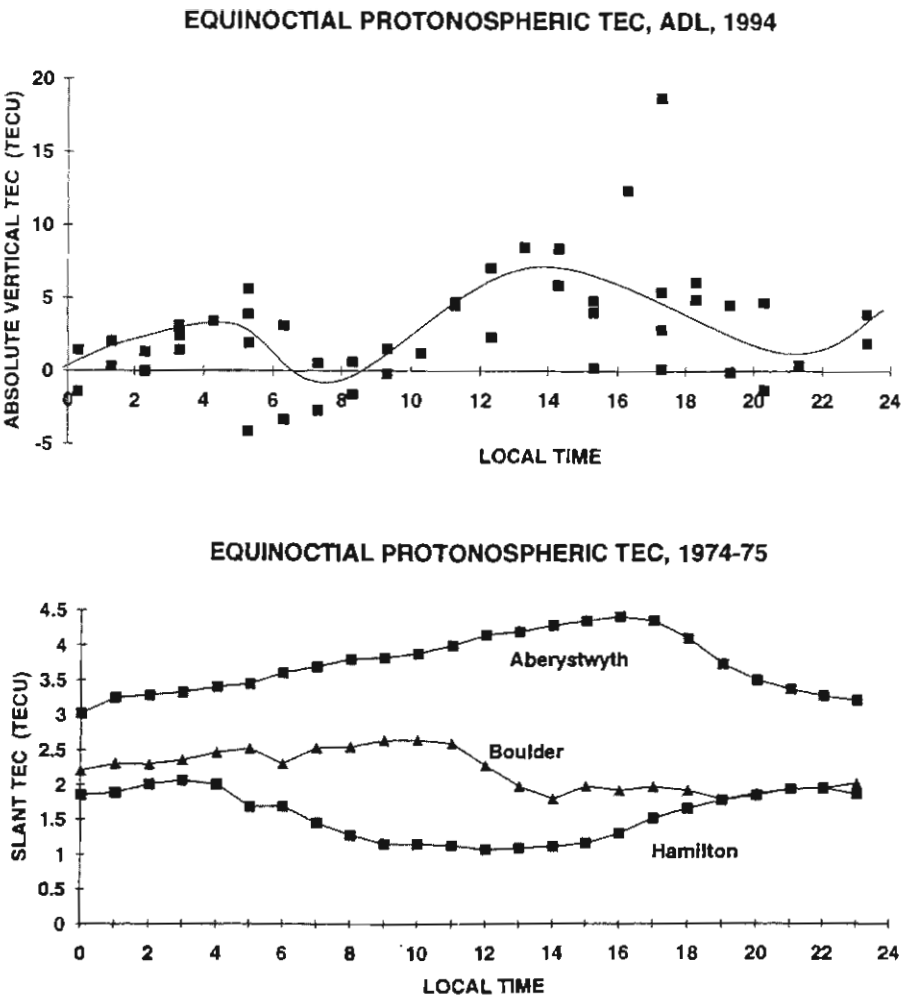


Figure 4. *Equinoctial electron content of the protonosphere from Salisbury in 1994, compared with measurements from three northern hemisphere sites.*

The Salisbury diurnal variation is similar to the variation of the slant protonospheric TEC observed in Aberystwyth but opposite to that observed in Hamilton. Kersley (1978) suggested that this difference between the Aberystwyth and Hamilton observations arises because the protonosphere at Aberystwyth has a conjugate point at a mid-latitude, whereas Hamilton has a conjugate point in an auroral latitude where the electron content is much lower than at mid-latitudes. The similarity with Aberystwyth could arise because Adelaide observations also relate to a conjugate point at a middle geographic latitude, namely at latitude 52.8°N, longitude 144.2°E, in the region north of Japan, just off the east coast of Russia.

Protonospheric TEC may also be determined from GPS TEC by deducting model ionospheric TEC (IRI and/or PIM). Another approach would be to deduct from GPS TEC the Navy Navigation Satellite System (NNSS) TEC measurements. NNSS TEC refers to a height up to 1000 km, and so a 'correction' of approximately 2 TECU (corresponding to TEC from 1000 to 2000 km) must be deducted from the protonospheric TEC calculated in this way.

The electron content of the protonosphere does not change much diurnally, but it is significantly depleted during major magnetic storms, and can take several days to recover to pre-storm values. For this reason in 1991–92, shortly after solar maximum, magnetic storms led to reduced protonospheric TEC compared with 1994–95, approaching solar minimum, as seen in Table 3 in which midday and midnight data are combined. Although Table 3 was determined for only a limited number of days, it does support the suggestion that in 1991–93 the GPS measurements would contain a smaller component from the protonosphere. From this observation, coupled with the occurrence of larger values of ionospheric TEC near solar maximum, it is concluded that in 1991–93 GPS TEC would be approximately the same as ionospheric TEC.

18.4 Ionospheric slab thickness

Slab thickness may be regarded as the depth of an imaginary ionosphere which has the same TEC as the actual ionosphere and uniform electron density equal to the maximum electron density of the actual ionosphere.

By definition,

$$\text{slab thickness} = \text{ionospheric TEC} / N_mF_2$$

where N_mF_2 is the peak F_2 region electron density (electrons m^{-3}), and TEC is the total electron content (electrons m^{-2}) of a vertical column of 1 m^2 cross-section up to the 'top' of the ionosphere, taken for convenience to be 2000 km.

Slab thickness includes information on both the topside and bottomside ionosphere, and indicates the electron density versus height profile; for instance, the 'sharper' the peak electron density, the smaller is the slab thickness.

The values of N_mF_2 which were used to compute slab thickness were calculated using the following relationship (e.g. Goodwin *et al.*, 1995a):

$$N_mF_2 = 1.24 \times 10^{10} (f_oF_2)^2 \qquad \text{(Davies and Liu, 1991)}$$

where f_oF_2, the F_2 region ordinary critical frequency in MHz, is measured from ionograms recorded at five minute intervals at Salisbury. Values of slab thickness were determined at Salisbury from July 1991 to February 1995 for three to five days per month. This period was during the descending phase of solar cycle 22 which had a broad peak from

mid-1989 to late 1991, and reached a minimum in May 1996. The present paper relates to earlier work (Goodwin *et al.*, 1995a, 1995b; Breed *et al.*, 1995; Breed, 1996).

Slab thickness values determined from GPS TEC include the electron contents of both the ionosphere (up to 2000 km) and most of the protonosphere (from 2000 km to 20 000 km). In ionospheric studies, slab thickness would more appropriately be defined as the slab thickness of the F_2 region only, ie:

$$\text{slab thickness } (F_2) = \text{TEC}(F_2)/N_mF_2$$

where $\text{TEC}(F_2)$ is the total electron content of the ionospheric F_2 region, that is GPS TEC from which had been deducted from the protonospheric TEC, (as well as, desirably, the daytime E and F_1 region TEC).

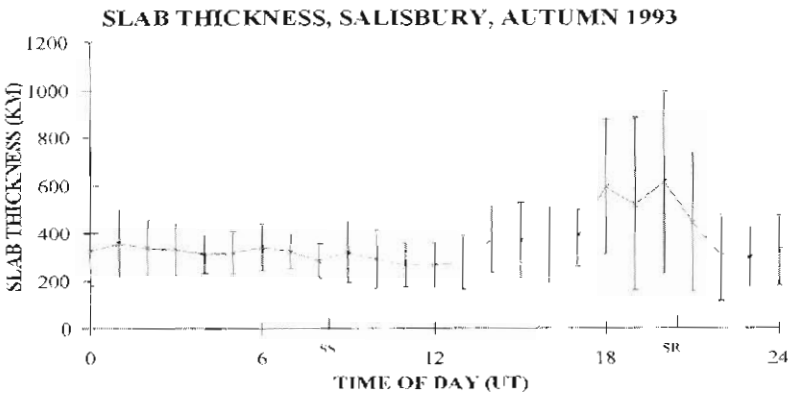
In the present paper, slab thickness is determined from GPS TEC recorded in 1991–93 near solar maximum when the protonospheric component was minimal. This calculated slab thickness is therefore taken to be approximately equal to ionospheric slab thickness.

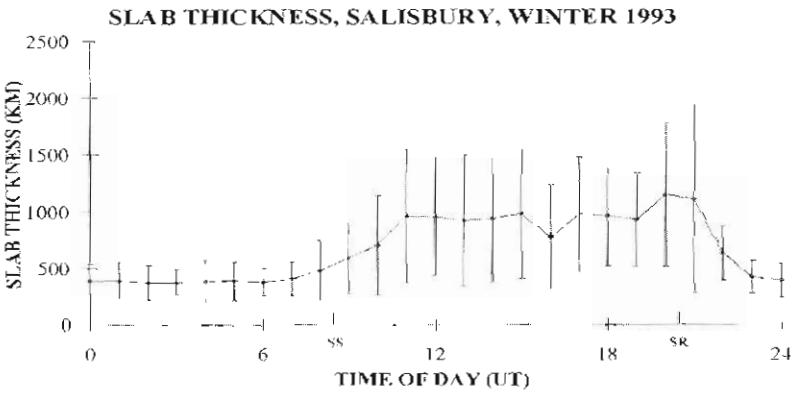
Figure 5 shows diurnal plots of slab thickness in different seasons in 1993 for Salisbury. (In Figures 5–8, seasons are: summer: November-February; winter: May–August; equinox: March, April, September, October; autumn: March, April; spring: September, October). TEC observations were made on 3–5 days each month. For Salisbury UT is 9.5 hours behind local time (Central Standard Time, CST); 2030 UT is around sunrise (SR) and 0830 UT is around sunset (SS). A pre-dawn peak in the slab thickness at 1800 to 2100 UT (0330–0630 CST) may be observed in these plots. The pre-dawn peak was also found in the seasonal plots from 1991 to 1995, except during summer in 1991–92 and 1992–93 when the solar influence was maximal.

The pre-dawn peak is explained in terms of the lowering of the F-layer height just before sunrise into a region of greater neutral air density. This causes a corresponding increase in the loss rate of ionisation due to recombination especially in the lower F_2 region which includes the electron density peak, N_{max} (Titheridge, 1973). As a result, there is a faster decrease in N_{max} compared with the decrease in TEC, since approximately 70% of the total electron content occurs above the peak in the topside ionosphere, where the loss rate is lower. Thus, an enhancement occurs in the slab thickness, (TEC/N_{max}).

Figure 5 indicates that the slab thickness is greater at night than in the day, and at midday (0230 UT) it is greater in summer than in winter. These observations agree with those of earlier workers in Australasia (Titheridge, 1973; Essex, 1978; McNamara, 1982) who used Faraday rotation measurements of TEC.

The winter night-time enhancement in slab thickness is due to a lowering (Titheridge, 1973) of the O^+/H^+ transition height (between the ionosphere and the protonosphere) at a time when solar influence





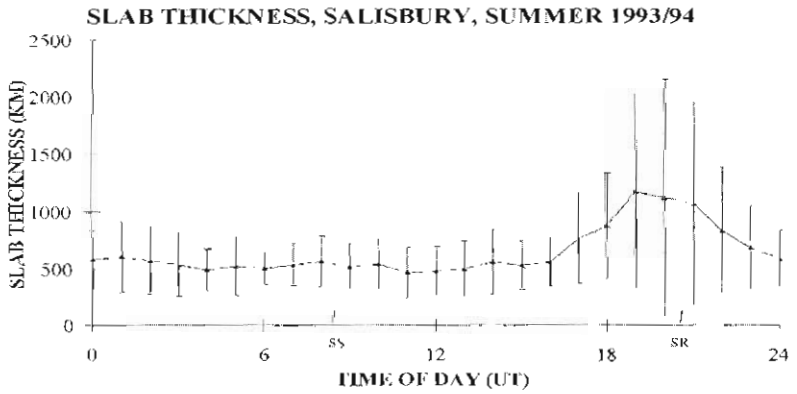


Figure 5. *Hourly averages of slab thickness calculated seasonally for 1993 are shown.*

is minimal. This transition height can descend to as low as 420 km near midnight in winter. The scale height (inversely proportional to atmospheric molecular weight) is greater above the O^+/H^+ transition height. It follows then that the further down the penetration of the O^+/H^+ transition height into the ionosphere, the greater the average ionospheric scale height, and hence the greater the slab thickness (which is directly proportional to scale height).

During 1991 to 1993, near solar maximum, GPS TEC had a comparatively small protonospheric component, and so slab thickness derived from GPS TEC was then a more reliable measure of the ionosphere than at other times in the solar cycle. Figure 6 shows latitudinal plots of slab thickness for summer (midday and midnight, November to February) in 1991–92 and 1992–93. Summer slab thickness is seen to be approximately constant (versus latitude) around 500 km at midday and around 300 km at midnight, with no substantial latitudinal change. In Figure 7 similar plots are shown for the equinoctial months (March, April, September and October) in 1991–92 and 1992–93. Again slab thickness is approximately constant against latitude with values of around 500 km at midday 1991–92, 300 km at midday 1992–93, 300 km at midnight 1991–92 and 400 km at midnight 1992–93. Figure 8 shows similar plots for winter (May to August) 1991 and 1992 for which slab thickness is approximately constant at around 250 km at midday and around 400 km at midnight. Once more there is no substantial latitudinal change. These observations, which indicate an absence of latitudinal dependence of slab thickness in any season, are in agreement with the conclusions of Davies and Liu (1991) and Liu and Davies (1990). The present observations also indicate that near solar maximum and in summer, midday slab thickness is greatest, whereas midnight slab thickness is least.

Finally, in Figure 9, there is a good agreement between two days' slab thickness determinations near solar maximum in October 1991 and March 1992, compared with slab thickness calculated from the IRI and PIM models. Slab thickness values of the order of 350 km were both observed and predicted.

18.5 Discussion and conclusions

The present measurements of GPS TEC, which are centred around Salisbury at 35°S latitude, are the only southern hemisphere TEC data extended in time, from 1991 onwards. In particular, midday TEC is seen to decrease with declining solar activity.

The determination of the protonospheric TEC component (from approximately 2000 km to 20 000 km) is exemplified in the present paper by deducting Faraday rotation measurements of TEC from GPS TEC.

In 1991–93, near sunspot maximum, when protonospheric TEC is negligibly small, slab thickness derived from GPS TEC is approximately

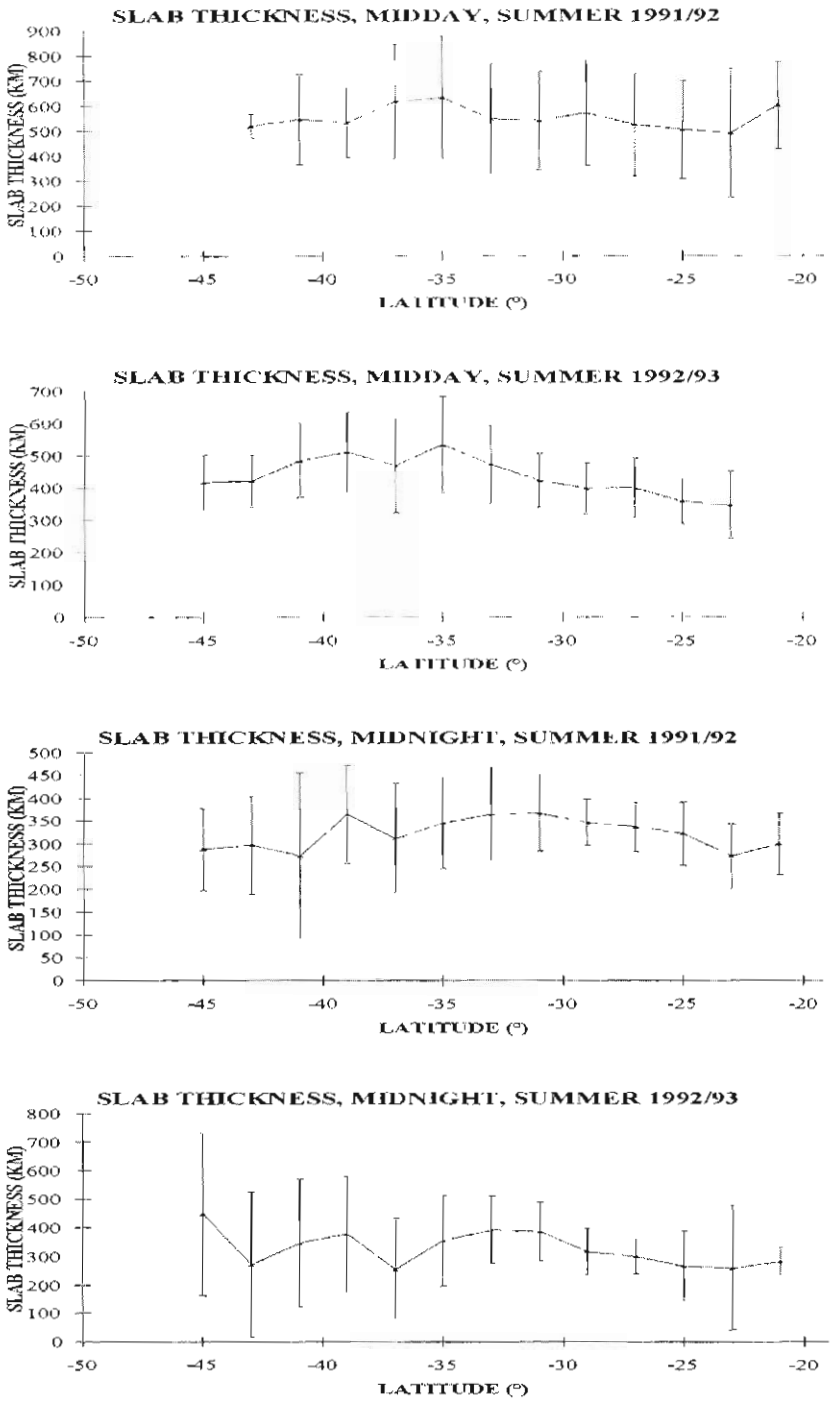


Figure 6. *Latitudinal variations in slab thickness measured from Salisbury for summer 1991–92 and 1992–93 are shown.*

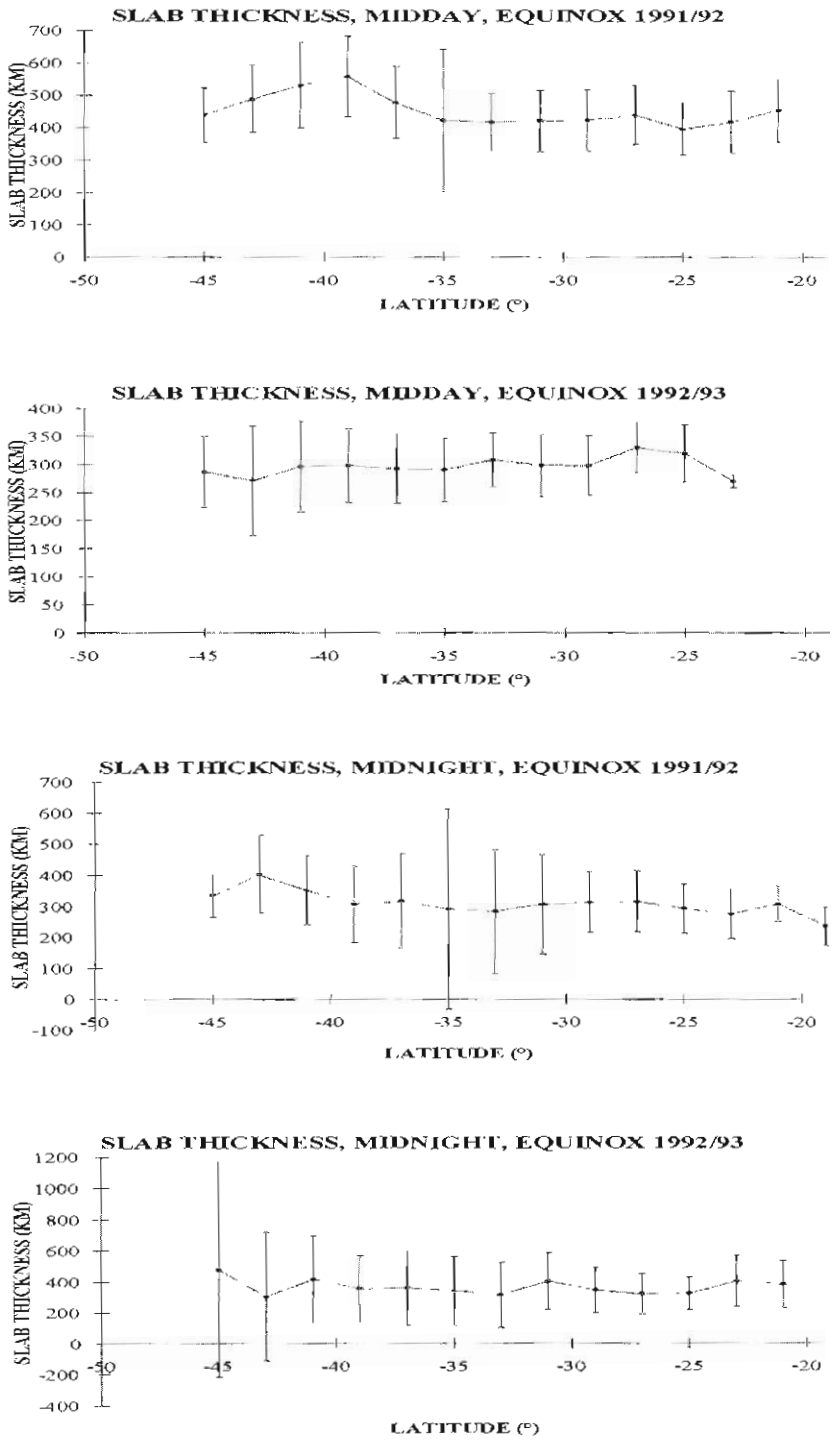


Figure 7. *Latitudinal variations in slab thickness measured from Salisbury for equinox 1991–92 and 1992–93 are shown.*

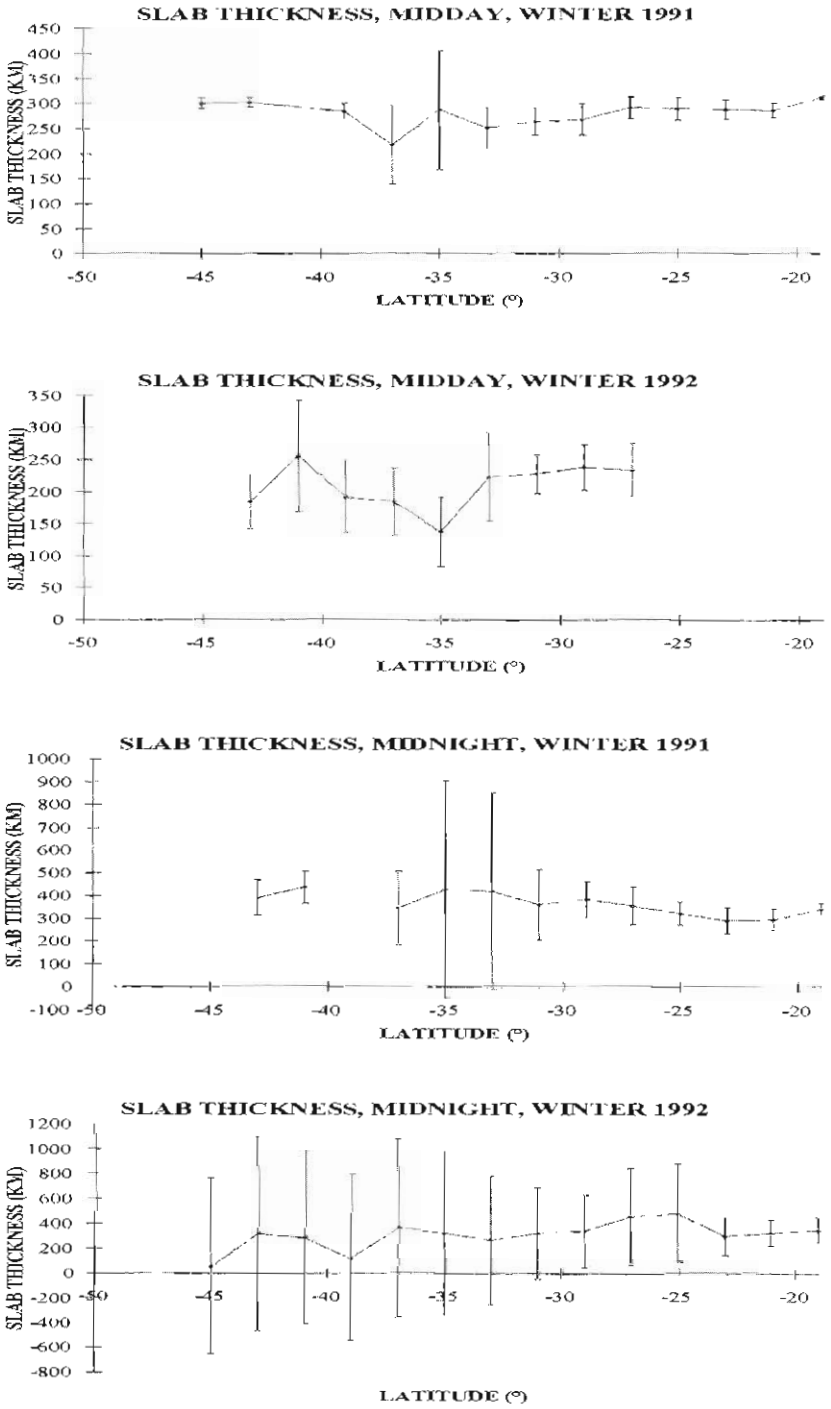


Figure 8. *Latitudinal variations in slab thickness measured from Salisbury for years 1991–92 are shown in winter.*

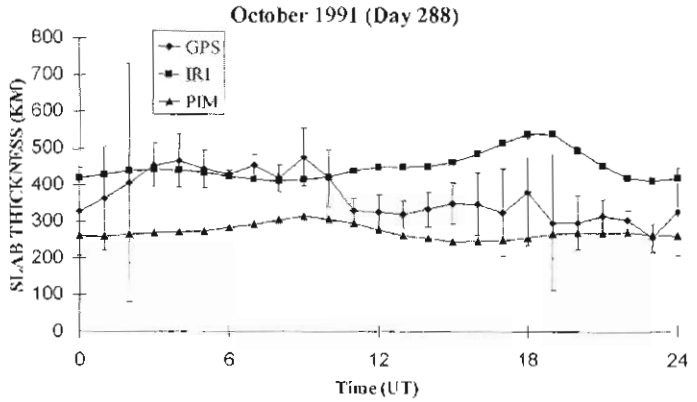


Figure 9(a). *A comparison of diurnal slab thickness values from GPS measurements and predictions from the IRI and PIM ionospheric models for D288, October 1991.*

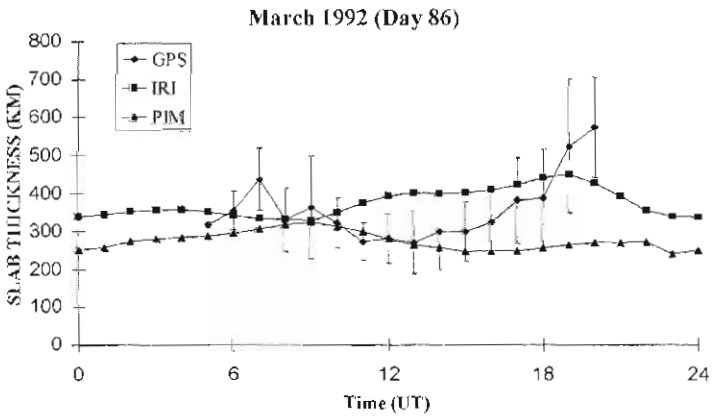


Figure 9(b). *A comparison of diurnal slab thickness values from GPS measurements and predictions from the IRI and PIM ionospheric models for D086, March 1992.*

the same as ionospheric slab thickness. In short, at times near solar maximum, GPS TEC and slab thickness derived from it, can be taken to represent the ionosphere. In any season and year in the declining solar cycle, slab thickness was observed to be substantially constant independent of latitude. Midday slab thickness is greater in summer than in winter, and greater near sunspot maximum than near minimum.

18.6 Acknowledgments

Sincere thanks are expressed to the late John Silby and to Alan Padgham of the Microwave Radar Division, DSTO, for painstakingly providing GPS receiver data from Salisbury, and to Ken Lynn of the Radio Wave Propagation Group, DSTO for the use of Salisbury ionograms.

References

Almeida, O.G., Garriott, O.K. and Da Rosa, A.V. (1970). Determination of the columnar electron content and the layer shape factor of the plasmasphere up to the plasmapause. *Planetary and Space Science*, 18: 159–170.

Bilitza, D. (1990). International Reference Ionosphere 1990, URSI/COSPAR Task Group on the IRI, NSSDC/WDC-A-R&S 90–22. National Space Science Data Center/World Data Center A for Rockets and Satellites, Greenbelt, Maryland, USA, November.

Breed, A.M. (1996). Investigations of the ionosphere over Australia using satellite transmissions. Ph.D. Thesis. School of Applied Physics, University of South Australia.

Breed, A.M., Vandenberg, A.M., Goodwin, G.L., Silby, J.H. and Essex, E.A. (1995). Measurement of the protonosphere in improving position determination using GPS satellites. In: Skellern, D. and Steel, J. (Eds.). *Workshop on Applications of Radio Science (WARS 95) Digest of Papers*. ISBN No. 1 86408 0302. Australian Committee for Radio Science, Canberra.

Daniell, R.E., Brown, L.D., Anderson, D.N., Fox, M.W., Doherty, P.H., Decker, D.T., Sojka, J.J. and Schunk, R.W. (1995). Paramaterized ionospheric model: a global ionospheric paramaterization based on first principle models. *Radio Science*, 30: 1499–1510.

Davies, K., Fritz, R.B. and Gray, T.B. (1976). Measurements of the columnar electron contents of the ionosphere and plasmasphere. *Journal of Geophysical Research*, 81: 2825–2834.

Davies, K. and Liu, X.M. (1991). Ionospheric slab thickness in middle and low latitudes. *Radio Science*, 26: 997–1005.

Essex, E.A. (1978). Ionospheric Total Electron Content at southern mid-latitudes during 1971–74. *Journal of Atmospheric and Terrestrial Physics*, 40: 1019–1024.

Goodwin, G.L. and Breed, A.M. (1999). Solar cycle variations in GPS observations of TEC in southern Australia. Paper presented at Workshop on Ionospheric/Atmospheric Effects on Space Based Assets. Defence Science and Technology Organisation, Salisbury, South Australia, January 1999.

Goodwin, G.L. , Silby, J.H., Lynn, K.J.W., Breed, A.M. and Essex, E.A. (1995a). Ionospheric slab thickness measurements using GPS satellites in southern Australia. *Advances in Space Research*, 15: 125–135.

Goodwin, G.L., Silby, J.H., Lynn, K.J.W., Breed, A.M. and Essex, E.A. (1995b). GPS satellite measurements: ionospheric slab thickness and total electron content. *Journal of Atmospheric and Terrestrial Physics*, 57: 1723–1732.

Huang, Y-N. (1978). Solar cycle variation in the total electron content at Sagamore Hill. *Journal of Atmospheric and Terrestrial Physics*, 40: 733–739.

Kersley, L. and Klobucher, J.A. (1978). Comparison of protonospheric electron content measurements from the American and European sectors. *Geophysical Research Letters*, 5: 123–126.

Liu, X-M. and Davies, K. (1990). Global measurements of the slab thickness of the ionosphere. *Proceedings of the International Beacon Satellite Symposium*. Tucuman, Argentina. Pp. 159–163.

McNamara, L.F. and Smith, D.H. (1982). Total electron content of the ionosphere at 31°S, 1967–1974. *Journal of Atmospheric and Terrestrial Physics*, 44: 227–239.

Titheridge, J.E. (1973). The slab thickness of the mid-latitude ionosphere. *Planetary and Space Science*, 21: 1775–1793.

19. AN INVESTIGATION OF SOUTHERN HIGH-LATITUDE IONOSPHERIC IRREGULARITIES USING TOTAL ELECTRON CONTENT MEASUREMENTS

B.S. Tate and E.A. Essex

Cooperative Research Centre for Satellite Systems
Department of Physics
La Trobe University
Bundoora Victoria 3083 Australia
(email: b.tate@latrobe.edu.au)

Abstract

During the period 1993–96, transmissions from the US Navy Navigational Satellite System (NNSS) polar orbiting satellites, received at Casey station (66.28°S, 110.5°E geographic) in Antarctica, were used to investigate the amplitude scintillations at 150 MHz and the Total Electron Content (TEC) obtained from the differential phase technique. Investigations of ionospheric polar patches, and polar holes were carried out. The phase and amplitude data were recorded using a JMR-1 satellite receiver system, with all measurements logged on a PC. In order to locate the patches and polar holes, the TEC data were mapped in magnetic local time (MLT) versus invariant latitude (Λ) magnetic coordinates. Patches were observed in the polar cap region, at various locations and times during the April–August period of the years 1994 and 1995. The polar hole, a feature of the 0000–0300 MLT region at -75° to -80° Λ, was also observed at various times through the winter months. Average monthly TEC values were calculated, and compared to those obtained from the Parameterised Ionospheric Model (PIM) version 1.4, Daniell *et al.* (1995).

19.1 Introduction

The polar cap station Casey is the ideal location to investigate the occurrence of ionospheric holes, patches and associated radio scintillations in the southern high-latitude ionosphere. TEC was derived using measurements from the constellation of NNSS polar orbiting satellites. The NNSS satellite constellation consists of four active satellites transmitting on 150 and 400 MHz, with several others on standby at any given time. Completing approximately thirteen orbits per day, each satellite is visible two or three times a day above an elevation of 45° at any given fixed site.

TEC is defined as the number of electrons in a volume of one square meter in cross-section, extending along a ray path from a satellite to a ground receiver.

$$TEC = \int_p N(s)ds \tag{1}$$

where, N(s) is the electron density.

In terms of the differential phase experiment, Equation (1) can be modified to:

$$(Vertical)TEC = \int_0^h Ndh = A\cos\chi_m\left(\int_{t_1}^{t_2} Fdt + D\right) \tag{2}$$

where, A is a constant = 0.02056 × 10^{15} (SI units), and F is the beat frequency observed between two frequencies f and mf.

Vertical TEC is obtained using the integration constant D, which is derived from ionosonde obtained f_0F_2 data as shown in Equation (3):

$$(Vertical)TEC = 1.24x10^{13}\left(f_o f_2\right)^2\tau \tag{3}$$

where τ is the ionospheric slab thickness, in km.

A linear plot of vertical TEC calculated from Equation (2) versus vertical TEC calculated from Equation (3) intersects the 'Equation (2)' axis at a point, A x D, from which D can be determined. The value of D can then be inserted into Equation (3) to obtain absolute vertical TEC values. A complete derivation of the above equations can be found in Breed (1992).

The TEC values are plotted using magnetic co-ordinates (MLT versus Λ) and geographic co-ordinates (UT versus geographic latitude).

Variations in amplitude, phase or angle-of-arrival of radio signals are known as scintillations. These effects result when a radio wave propagating through the ionosphere is diffracted by irregularities that are moving relative to the ray path. Amplitude scintillations are characterised by the scintillation index S_4 (defined in Equation (4)), which provides a 'level of scintillation activity'.

$$S_4 = \sqrt{\frac{\left(\overline{P^2} - \overline{P}^2\right)}{\overline{P}^2}} \tag{4}$$

where, P is the received power and $\overline{P}$ is the mean received power.

19.2 Experiment

Casey station (formerly Wilkes) was the third permanent ANARE station to be established on the Antarctic mainland. Casey is situated at the geographic location of 66.3°S, 110.5°E, and at the invariant latitude of -80.4°. Its position, poleward of both the ionospheric cusp and auroral oval region, makes it an ideal location to study polar patches and to monitor the polar hole.

Phase and amplitude data from NNSS satellite signals were recorded using equipment located at Casey. The equipment consisted of a JMR satellite receiver, a 24 MHz amplitude receiver and an IBM-

compatible PC. The PC was used to control the JMR receiver and to record the differential phase and signal strengths to hard disk. High quality signals were obtained from the NNSS satellites due to the use of an antenna that consists of two, short helical coils tuned to 150 and 400 MHz respectively. These coils have polar patterns that enable high gain signals at high elevations to be received. The JMR receiver allows amplitude variations to be recorded, together with differential phase. Scintillation measurements were taken using a receiver, attached to the 24 MHz IF output of the JMR. The f_oF_2 data, used for TEC baseline calibration (Equation (3)) were collected using the DPS-4 ionospheric sounder located at Casey, provided courtesy of Atmospheric and Space Physics, Australian Antarctic Division. Data were collected for approximately 15 days per month, consisting of five days of scintillations and TEC at 50 Hz resolution and 10 days of TEC only at 10 Hz resolution being recorded. Monthly samples of both differential phase and scintillations were sent to La Trobe University via ftp once a month. The bulk of the data was archived and returned to Australia at the end of the year.

19.3 Results

19.3.1 Patches

Large scale plasma irregularities in the F-region ionosphere are produced near or equatorward of the dayside auroral zone and convect across the polar cap in the antisunward direction (Weber *et al.*, 1984). Plasma enters the polar cap as a tongue of ionisation which provides the source for enhanced polar cap F-region plasmas, or 'patches' which are observed at high polar latitudes propagating away from noon. These patches are localised regions of increased F-region density (up to 5 to 10 times above background level), extending up to 1000 km across. Patches have been found to occur most often in April and August, and less likely to occur in January and July (Rodger and Graham, 1996). Patches are most likely to occur when the B_z component of the interplanetary magnetic field (IMF) is southward, or when the planetary magnetic index $K_p > 4$. It is possible to track these patches using the method of differential phase determination of columnar TEC, and hence determine some of the dynamics of the polar cap ionosphere (Beggs *et al.*, 1994). An example of a TEC plot showing a polar patch on day 241, 1995, is seen in Figure 1. Also shown in Figure 1 are the associated amplitude variations and S_4 index recorded during the pass. For the patch observed on day 241 (shown in Figure 1), a 'burst' of strong scintillation is observed during the patch event. S_4 levels rise as high as 0.7 during this period.

Using the above conditions as a guide, patches were located on several occasions during the period of April 1994 through to August 1995. The patches from 1995 are plotted as MLT versus invariant latitude polar plots (see Figure 2). All patches shown were observed during periods of IMF $B_z < 0$. K_p levels varied from 0.3 (0^+) up to 7.0 (7). Approximately 50% of the patches occurred during periods with $K_p > 3.7$ (4^-). These results tend to agree with the accepted conditions for patch behaviour.

The f_0F_2 records were also analysed in an attempt to verify that the patches observed in TEC records could also be seen in f_0F_2 records. This could not be verified with any level of certainty, since the f_0F_2 records for the days in question contained many omissions of data. Amplitude scintillation measurements were taken using the method developed by Kersley *et al.* (1988) and were used to identify scintillation producing patches. Not all patches identified were observed to produce scintillations. The S_4 index was observed to vary over the course of a patch in all scintillation producing patches.

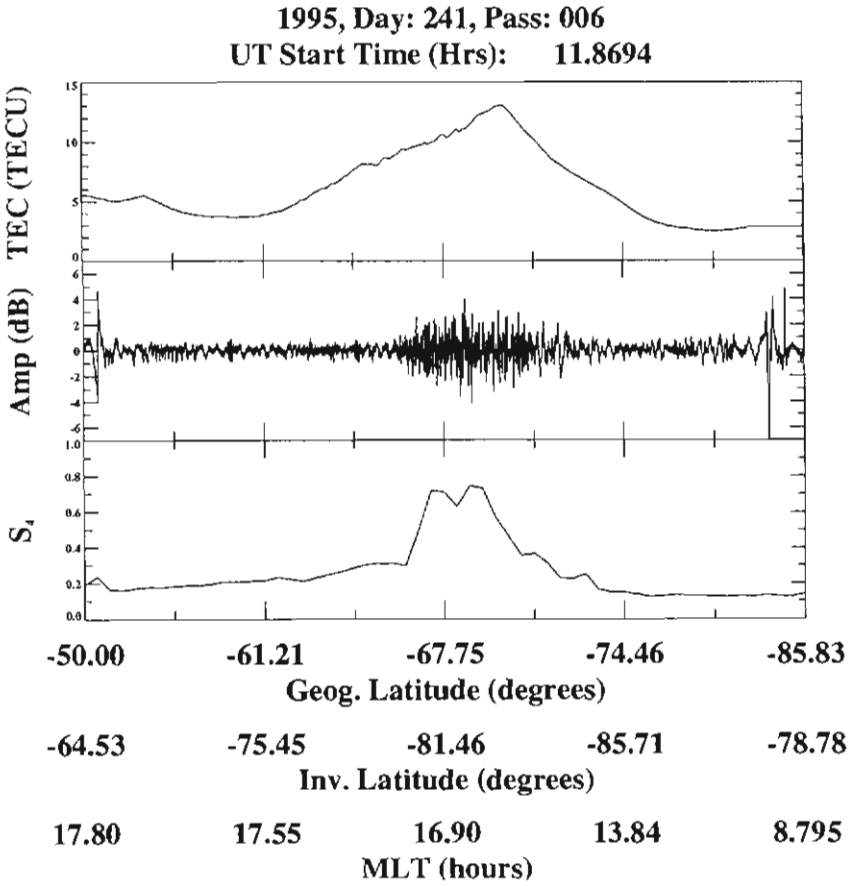


Figure 1. *Example of a polar patch seen in TEC data, with associated amplitude scintillation and S_4 index.*

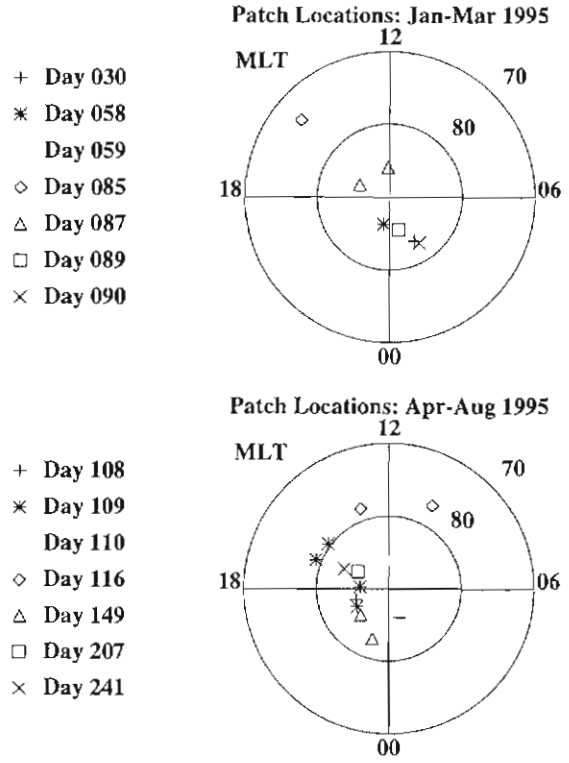


Figure 2. *Locations of patches for 1995, plotted in MLT versus invariant latitude (Λ).*

19.3.2 Holes

The ionisation 'hole' poleward of the nightside auroral zone first became evident during the late 1970's (Brinton *et al.*, 1978; Crowley *et al.*, 1993). The hole arises due to the long transport time of ionisation from the dayside across the dark polar cap, and from plasma stagnation (see Kelley, 1989 for a detailed description). The polar hole is expected to be located in the post-midnight sector (0000–0300 MLT) when the IMF B_z is southward with a positive B_y, whilst it is expected to be found in the pre-midnight sector of the polar cap (2100–2400 MLT) when the IMF B_z is southward with a negative B_y (Beggs *et al.*, 1994). It is also expected that the polar hole ionisation density is lowest during winter with a minimum density in the southern hemisphere at 2030 UT for all seasons. This hole can also be located, using differential phase TEC measurement in the 2100–2400 MLT region for IMF $B_y < 0$ and in the 0000–0300 MLT region for IMF $B_y > 0$. An example of a hole seen in TEC measurements on day 146, 1995, along with the associated amplitude variations and S_4 index, is shown in Figure 3. Scintillation levels slowly rise during the hole event as S_4 rises from 0.2 up to a peak of 0.4 in the 3.0 to 2.5 hours

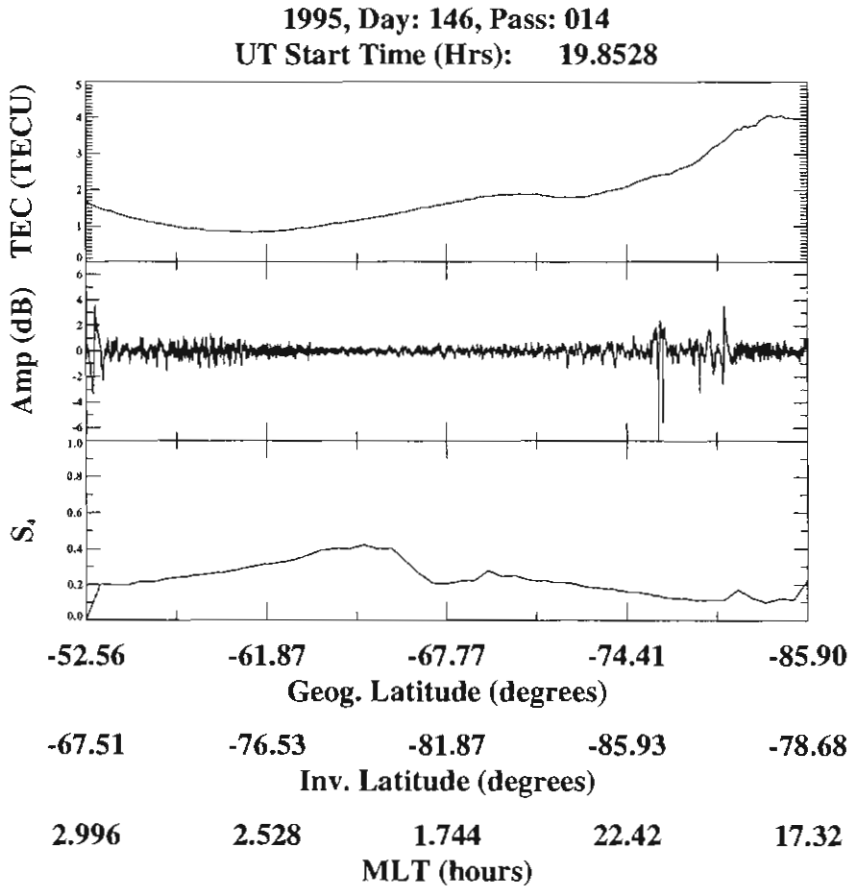


Figure 3. Polar hole as seen in TEC data, with associated amplitude scintillation and S_4 index.

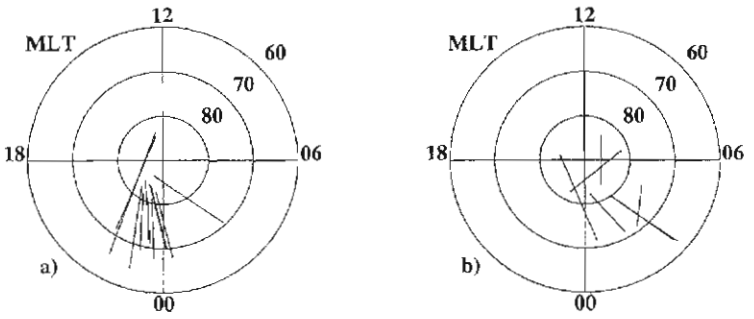


Figure 4. Locations of holes plotted in MLT versus invariant latitude (Λ) with; (a) IMF B_y negative, and (b) IMF B_y positive.

MLT region. A large proportion of passes showing polar holes exhibited similar scintillation behaviour. The hole was located at various times, during the period of April to August 1995, for both IMF B_z, $B_y < 0$, and for IMF $B_z < 0$, $B_y > 0$, as shown in Figure 4. As expected, the hole was only observed during conditions of southward IMF ($B_z < 0$), and seen to be situated in the -70° to -80° invariant latitude (Λ) region, for all IMF B_y. The hole was also seen to be located in the 2100–2400 MLT region for IMF $B_y < 0$ and in the 0000–0300 MLT region for IMF $B_y > 0$.

19.3.3 Comparisons with models

Selected TEC passes were compared to results from the Parameterised Ionospheric Model (PIM) version 1.4 (Daniell *et al.*, 1995). Figure 5 shows a comparison of experimental data, with theoretical data, both plotted in geographic co-ordinates. Also shown is the experimental data, plotted in magnetic co-ordinates. The results obtained from PIM show only a large-

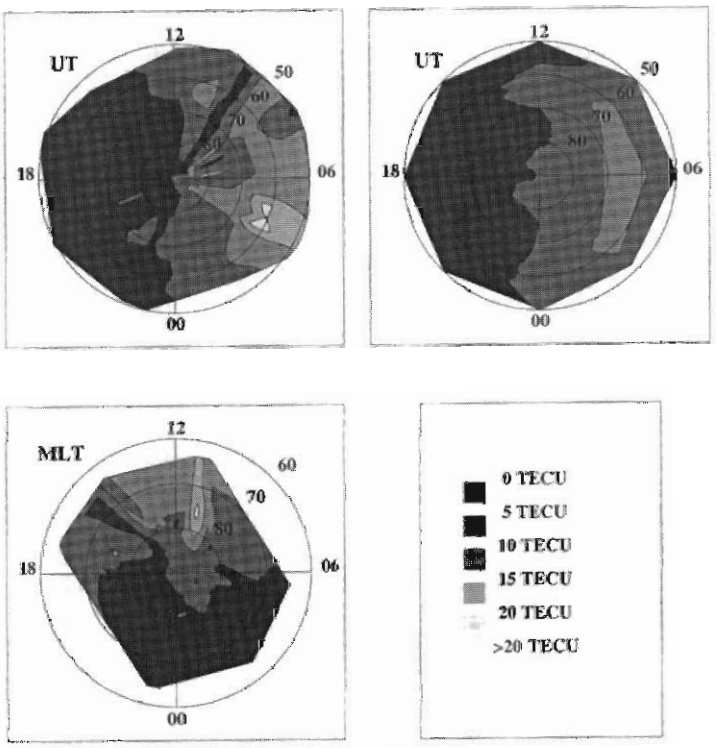


Figure 5. *Comparison of average TEC values, plotted in both MLT versus invariant latitude and UT versus geographic latitude, to PIM TEC plotted in UT versus geographic latitude for the month of April, 1995.*

scale correlation to the experimental data. PIM does not give the correct background levels of TEC for most of the passes analysed, although the monthly average values are close to that expected, if only on a large scale. The average background TEC levels predicted by PIM were accurate to within 2–3 TEC units (where one TEC unit is equal to 1×10^{16} electrons m^{-2}) to that recorded experimentally. PIM is not able to predict the position of polar holes or patches, as these are dependent on IMF parameters.

19.4 Discussion and conclusions

Patches were observed in the -80° to -90° invariant latitude region, predominantly between 1200–2400 MLT, on several days throughout the periods of April to November 1994, and April to August 1995. Patches were only observed during periods of southward interplanetary magnetic field (IMF $B_z < 0$) conditions. November 1994 and April 1995 were the times when patches were most often observed. TEC levels rose by a factor of 2 to 4 times that of background levels during the patch event. Strong scintillation was observed on several occasions when TEC levels rose significantly compared to background levels. The ionisation 'hole' was observed on several days during the months of April to August 1995. The hole was predominantly situated between -70° to -80° invariant latitude, between 2100–2400 MLT ($B_y < 0$), and 0000–0300 MLT ($B_y > 0$). As predicted, holes were found in two regions, dependent on whether IMF B_y was positive or negative. This study shows that the method of differential phase determination of total electron content can be used to locate patches and holes in the polar cap ionosphere.

19.5 Acknowledgements

Special thanks go to Australian Antarctic Division personnel Lloyd Symons, Dale Siver, Dr Didier Monselesan and Dr Darryn Schneider for assistance with data collection. Thanks also to Dr Anthony Breed from the Australian Antarctic Division for assistance with data analysis software. Also thanks to Mr Robert Polglase of La Trobe University for help with technical matters, and to Dr Paul Smith for providing f_oF_2 and IMF data. IMF data obtained from the WIND satellite, courtesy of the National Space Science Data Centre. This research is supported by an Antarctic Science Advisory Committee (ASAC) grant, and by logistic support from the Australian Antarctic Division.

References

Beggs, H.M., Essex, E.A. and Rasch, D. (1994). Antarctic polar cap total electron content observations from Casey Station. *Journal of Atmospheric and Terrestrial Physics*, 26: 659–666.

Breed, A.M. (1992). Ionospheric Total Electron Content studies using satellite radio transmissions. Master of Applied Science Thesis. University of South Australia.

Brinton, H.C., Grebowsky, J.M. and Brace, L.H. (1978). The high latitude F-region at 300 km: thermal plasma observations from AE-C. *Journal of Geophysical Research*, 83: 4767–4776.

Crowley, G., Carlson, H.C., Basu, S., Denig, W.F., Buchau, J. and Reinisch, B.W. (1993). The dynamic ionospheric polar hole. *Radio Science*, 28: 401–413.

Daniell, R.E. Jr., Brown, L.D., Anderson, D.N., Fox, M.W., Doherty, P.H., Decker, D.T., Sojka, J.J. and Schunk, R.W. (1995). Parameterized ionospheric model: a global ionospheric parameterization based on first principles models. *Radio Science*, 30: 1499–1510.

Kelley, M.C. (1989). The Earth's Ionosphere: Plasma Physics and Electrodynamics. San Diego Academic Press. Pp. 326–329.

Kersley, L., Pryse, S.E. and Wheadon, N.S. (1988). Amplitude and phase scintillation at high latitudes over northern Europe. *Radio Science*, 23: 320–330.

Rodger, A.S. and Graham, A.C. (1996). Diurnal and seasonal occurrence of polar patches. *Annales Geophysicales*, 14: 533–537.

Weber, E.J., Buchau, J., Moore, J.G., Sharber, J.R., Livingstone, R.C., Winningham, J.D. and Reinisch, B.W. (1984). F-layer ionization patches in the polar cap. *Journal of Geophysical Research*, 89: 1683–1694.

20. SCINTILLATION AND TEC STUDY OF THE HIGH-LATITUDE IONOSPHERE OVER CASEY STATION ANTARCTICA

N.M. Shilo[1], E.A. Essex[1] and A.M. Breed[2]

[1] Cooperative Research Center for Satellite Systems
School of Physics
Faculty of Science and Technology
La Trobe University
Bundoora Victoria 3088 Australia
(email: e.essex@latrobe.edu.au)

[1] Atmospheric and Space Physics
Australian Antarctic Division
Kingston Tasmania 7050 Australia

Abstract

During the summer season 1997–1998 a GPS Ionospheric Scintillation Monitor (GISM) was installed at Casey station, Antarctica, 66.28°S, 110.24°E, -80.38 invariant latitude. This system is capable of tracking up to 11 GPS satellites at the L1 frequency of 1575.42 MHz. The purpose of the GISM receiver is to automatically record scintillation parameters of amplitude and phase at a 50 Hz rate averaged over 60 s. These measurements will be used to map scintillation activity associated with the polar cusp, auroral oval and other ionospheric features such as patches. In conjunction with scintillation measurements, 1998 Total Electron Content (TEC) data were recorded by the Australian Surveying and Land Information Group (AUSLIG), using the L-band signals from GPS satellites for the Antarctic stations Casey, Davis, Mawson, and the sub-antarctic station Macquarie Island. From these data, 3–5 days are being selected for each month for Casey centered on the Regular World Days (RWD). These results are used to investigate the diurnal and seasonal variations in TEC, and other features such as patch occurrences.

20.1 Introduction

Time variations in phase, amplitude and angle-of-arrival of radio waves propagating through the ionospheric medium are known as ionospheric scintillations. Scintillations are due to the diffraction caused by irregularities in motion relative to the ray path. With solar activity due to reach its maximum in the year 2000, scintillation activity is expected to increase accordingly. This paper describes scintillation and TEC measurements made at Casey station, 66.28°S, 110.24°E, in Antarctica.

20.1.1 Amplitude scintillation

Amplitude scintillation is obtained by monitoring the index S_4. The S_4 index is derived from detrended signal intensity of signals received from satellites (Van Dierendonck *et al.*, 1993). The S_4 index, which includes the effects due to ambient noise, is defined as the normalised root mean square of the power P divided by the average power P as follows:

$$S_{4T} = \sqrt{\frac{\langle P^2 \rangle - \langle P \rangle^2}{\langle P \rangle^2}} \tag{1}$$

where $\langle\rangle$ represents the average values over a 60 s interval. Note: Equation (1) is referred to as the Total S_4.

20.1.2 Removing the effects of ambient noise

The Total S_4 defined in Equation (1) has a significant amount of ambient noise associated with it, which needs to be removed before further analysis. This is achieved by estimating the average signal-to-noise density (S/N_0) over a 60 s interval. This estimate is then used to determine the expected S_4 due to ambient noise (also known as S_4 correction) as follows:

$$S_{4N_0} \sqrt{\frac{100}{S/N_0} 1 + \frac{500}{19S/N_0}} \tag{2}$$

Replacing the S/N_0 with the 60 s estimate, $\overline{S/N_0}$, gives the S_4 due to noise. Hence, subtraction of Equation (1) and (2) yields Equation (3), which is the S_4 with the effect of ambient noise removed.

$$S_4 = \sqrt{\frac{\langle P^2 \rangle - \langle P \rangle^2}{\langle P \rangle^2} - \frac{100}{\overline{S/N_0}} 1 + \frac{500}{19\overline{S/N_0}}} \tag{3}$$

20.1.3 Phase scintillation

Phase measurements are obtained by monitoring the standard deviation, $\sigma\Delta\varphi$, and the power spectral density of the detrended carrier phase from signal received from satellites. The spectral slope is measured above 1 Hz from detrended carrier and the spectral strength is measured at 1 Hz from detrended carrier.

The detrending of the carrier phase is achieved by passing the raw 50 Hz phase measurements through a sixth-order high-pass Butterworth filter. This removes all the low frequency effects below its cut-off of 0.1 Hz (Van Dierendonck *et al.*, 1996).

The $\sigma\Delta\varphi$'s are computed over 1, 3, 10, 30 and 60 s intervals every 60 s using the 50 Hz detrended phase measurements. The 1, 3, 10, 30 and 60 s $\sigma\Delta\varphi$'s are further averaged over the 60 s interval.

20.2 Total electron content

The electron density is obtained by counting the number of electrons in a vertical column with a cross-sectional area of 1 m^2, extending from the GPS satellite to the observer. The electron density obtained is termed the Total Electron Content (TEC) (Klobuchar, 1991), and is mathematically expressed as:

$$TEC = \int_p N ds \tag{4}$$

where N is the electron density m^{-3} and p is the propagation path between the satellite and the detector. A TEC unit is defined as 1 x 10^{16} electrons m^{-2}. TEC measurements are generally derived from satellite radio signals observed at various angles, and are normally expressed as an equivalent vertical TEC by dividing the slant TEC by the secant of the elevation angle at a mean ionospheric height (i.e. 350–400 km). Here, vertical TEC is used, as it is more easily comparable than slant values at various angles.

20.3 Experimental set-up

The GISM was installed at Casey station and it consists of three major components (see Figure 1).

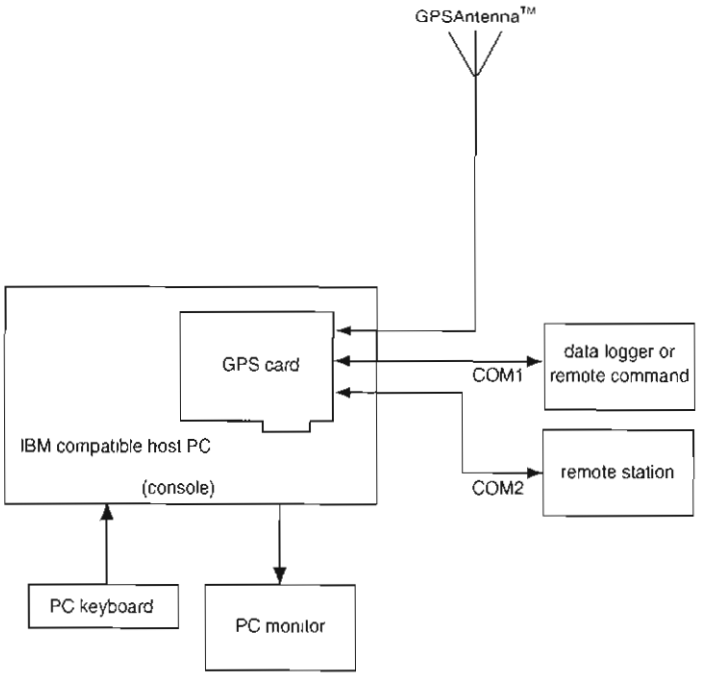


Figure 1. *This is an illustration of the operational configuration for the GISM. GPSCard™ PC. (Series Installation and Operational Manual page 21. Printed in Canada 1995 by NovAtel Ltd.)*

(i) GPS antenna (NovAtel's model 503 GPS antenna) mounted on a platform located on the roof of the Casey science building.

(ii) GPS receiver (NovAtel's 3951R GPSCard) installed in a computer located in the Atmospheric Space Physics (ASP) laboratory.

(iii) GSV-3003 5 or 10 MHz conversion board to provide an externally generated reference clock signal for the NovAtel GPSCard, generating the 20.473 MHz reference signal required by the GPSCard. The signal is phase locked to either the on board 10 MHz OCXO, or an externally applied 5 or 10 MHz local station frequency.

The receiver (GPSCard) is capable of tracking up to 11 GPS C/A signals at the L1 frequency, and measures amplitude and phase at a 50 Hz rate, and code/carrier divergence at a 1Hz rate for each satellite tracked. Menu-based programs GSV4000.EXE and ISMVIEW.EXE run the system. The first consists of three sub-menus: 'log selection', allowing the selection of binary data logs to be recorded onto the PC hard disk. The bandwidth of the 6th order Butterworth filters can be modified for both the phase high-pass filter (HPF), and the amplitude low-pass filter (LPF) used to detrend the raw phase and amplitude 50 Hz data respectively. This bandwidth can be varied between 0.001 and 5.000 Hz, and is set to 0.1 Hz. For the purpose of this campaign, the channel assignment had all 11 channels operational. The GISM is capable of monitoring the S_4, $\sigma\Delta\varphi$ and power spectral parameters of scintillation activity.

'Screen parameter selection' displays four parameters on the PC screen. It cannot accommodate any more. However, selections can be changed at anytime. The last menu, the 'display mode', enables viewing of data in real time. This option was not available for use during this campaign.

The raw data is stored in logging files known as SIN and DIV, and the processed data in ISMR. The SIN data log contains the raw 50 Hz rate phase and amplitude measurements, and the DIV data log records raw code/carrier divergence (C/No), due to the ionosphere. The average and standard deviation of the C/No are computed every minute. Signal-to-noise ratio (S/No) is also recorded every second.

The ISMR data log collects the reduced raw measurements from SIN and DIV every minute, on the minute, along with some other useful parameters. After processing, the raw data can be saved, however, for this campaign only the processed ISMR data is saved and used.

The ISMVIEW.EXE program permits the extraction of the ISMR data logs, for specified satellites, into individual files for further processing.

20.4 Preliminary results

The GISM has been running continuously for three years. The processed data is collected monthly via the use of the ISMVIEW program. As previously mentioned the receiver records the parameters S_{4T}, S_{4No} and S_4. A sample of data is illustrated in Figure 2 for September 1998, day 266, satellite PRN 18. The S_4 index here shows that there still appears to be some noise on the edges of the two passes, possibly due to a multi-

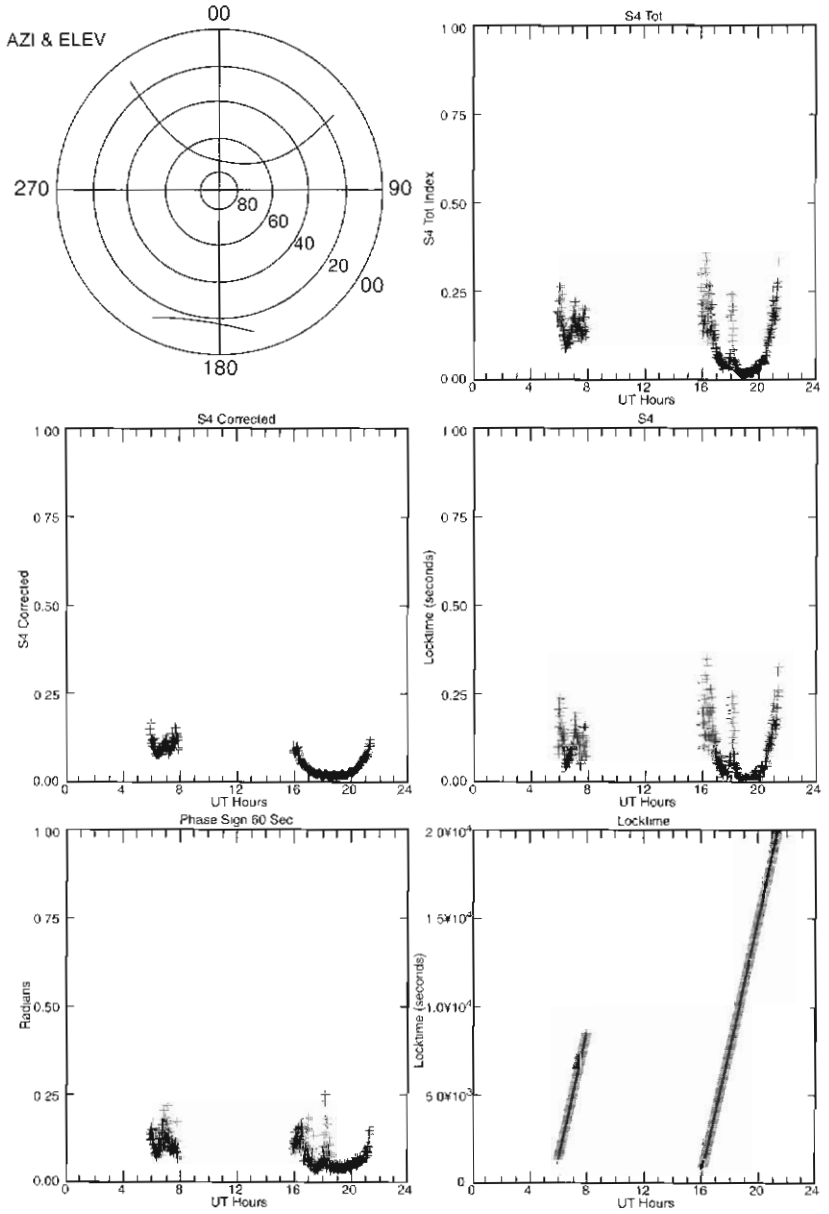


Figure 2. *Plots for satellite PRN 18, day 266 in UT time of the S_4 total, S_4 corrected, S_4 (with the effects of ambient noise removed), and 60 $\sigma\Delta\varphi$. The last plot shows the lock time of the satellite signal.*

path. There is, however, a prominent peak at the center of the second pass around 1800 UT hours, with S_4 measuring at 0.25 which could be indicative of some ionospheric disturbance (i.e. auroral activity or patches). The true source of this increase is yet to be determined. The 60 s sigma data (phase scintillation) for the same satellite also shows some scintillation activity around the same time.

The 'lock time' of the satellite signal is a useful parameter to have access to, for the simple reason that it allows the lock time of the receiver to be observed for each satellite pass. This helps in the initial observation of the noise content of the pass. If the receiver loses lock, or does not maintain lock long enough (i.e. < 240 s), then the data can be disregarded. This is the time it takes for the detrending high-pass-filter to re-initialise lock of the carrier phase signal. Looking at the lock time for PRN 18 in Figure 2, it can be seen that the receiver remains continuously locked on during this tracking period. Therefore, the increase that is observed around 1800 UT must be associated with some sort of disturbance activity.

The corresponding TEC for this scintillation data is represented in Figure 3. It too shows disturbed conditions around 1800 UT hours when the Interplanetary Magnetic Field (IMF) is southward, $B_Z < 0$. The TEC observed in Figure 3 may be related to gravity waves or patch activity (Beggs *et al.*, 1994; Huang *et al.*, 1998).

Patches are known as regions of enhanced ionisation, and are observed to drift antisunward across the polar cap when the IMF is directed southward ($B_Z < 0$). Not only do they occur when the IMF B_Z component is negative, but also during moderately disturbed conditions ($K_p > 4$) (Weber *et al.*, 1984), and during quiet periods (Basu *et al.*, 1990). The seasonal occurrence of patches has been observed to have a peak occurrence in April with a pronounced minimum in July, a maximum in

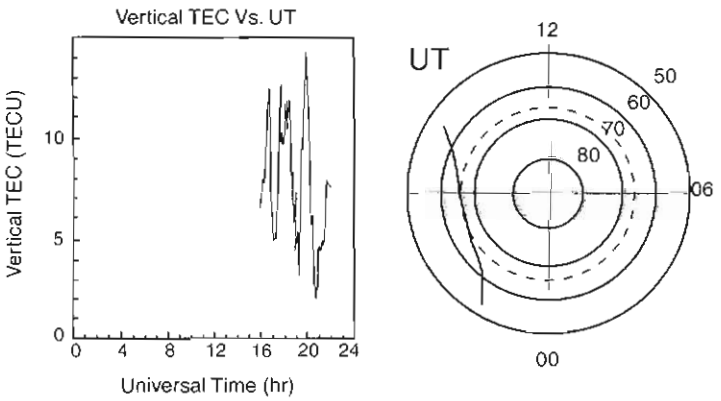


Figure 3. *Plot of the Vertical TEC day 266 and satellite pass for PRN 18. The dotted line in the polar plot represents the location of Casey Station*

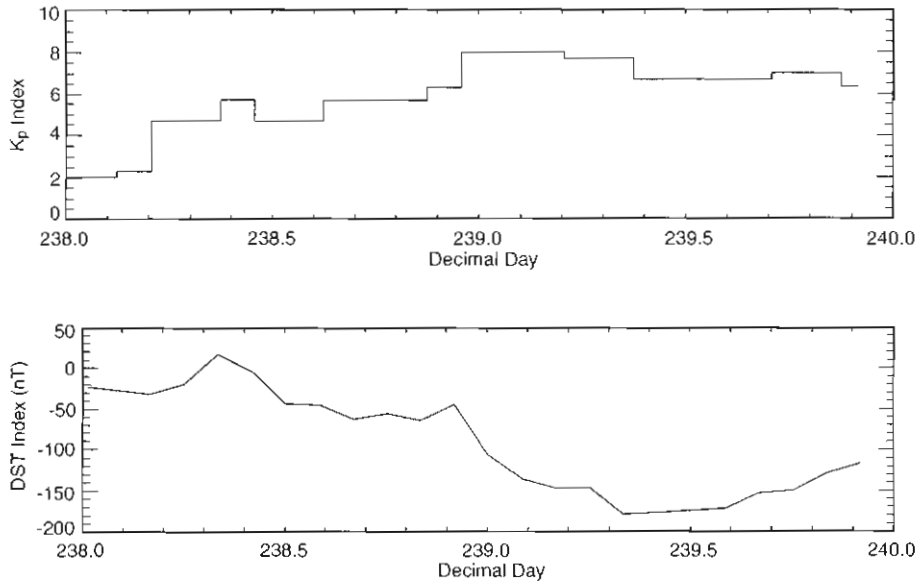


Figure 4. *Top plot shows the K_p index and the second plot the DST for days 238–239 of 1998. Data was obtained from: http://nssdc.gsfc.nasa.gov/omniweb/.*

August and minimum occurrence near mid-summer (December–January) (Tate *et al.*, 1998). Patches of smaller size and steep density gradients at their edges are associated with intense scintillation and interfere with trans-ionospheric radio signals. Based on previous studies increases in ionisation are defined to be at least 3 TEC units or more.

During late August 1998 magnetic storm conditions were observed for days 238 and 239. The K_p and Disturbance Storm Time (DST) are illustrated in Figure 4. A steady decrease in DST was observed towards the end of day 238, reaching a maximum decrease of -188 around 0900 UT on the following day, with a corresponding K_p value of 7.7. The maximum K_p observed for day 239 was 8.0. Figure 5 (day 238) shows a plot of vertical TEC for PRN 14 along with the satellite passes in both UT and MLT, to show their positions relative to both the geomagnetic and magnetic poles. This particular sample of TEC data was taken at the beginning of the storm activity and clearly shows an enhancement in the ionisation at 1700 UT of 10 TEC units with a K_p value of 5.7, possibly associated with a patch occurrence. The K_p values continue to rise during day 239 with numerous TEC enhancements observed as shown in Figure 6 for PRN 6. A noticeable increase can be seen during 0700–1000 UT with a peak in TEC of 6 TEC units at 0900 UT. IMF data is needed to determine if these observed increases are associated with patches. Unfortunately, IMF data was not available for these days.

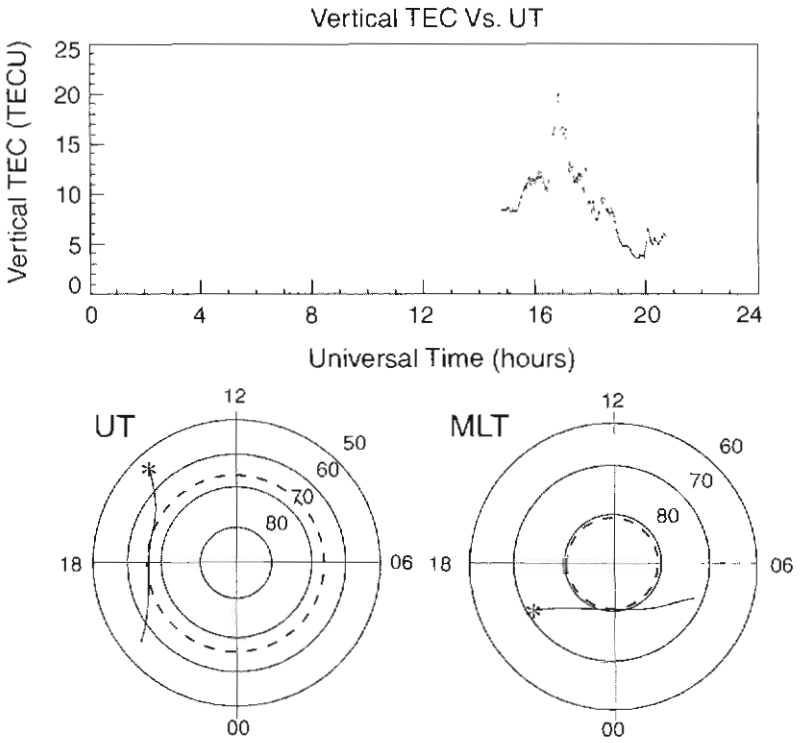


Figure 5. *Shows Vertical TEC and polar plots of passes for satellite 14 in UT/MLT, for day 238. The * in the polar plots denotes the start of the pass.*

20.5 Conclusion

With the approaching solar maximum, scintillation effects on GPS are becoming more frequent, with GPS receivers losing lock on satellites. At this stage however, scintillation levels observed are no higher than 0.25 for the high-latitude region. So far the receiver has performed well in keeping locked-onto a tracked satellite, although for a few cases it loses lock, at the beginning and end, of a satellite pass at low elevation angles. The cause of the scintillation observed in Figure 2 is still to be determined. The magnetic storm occurrence for August shows some increase in the ionisation over the two-day period. These enhancements have been observed to range anywhere form 6–10 TEC units with K_p values ranging from 5.7 to 8.0. Further analysis is needed to verify the cause of high-latitude enhancements in ionisation.

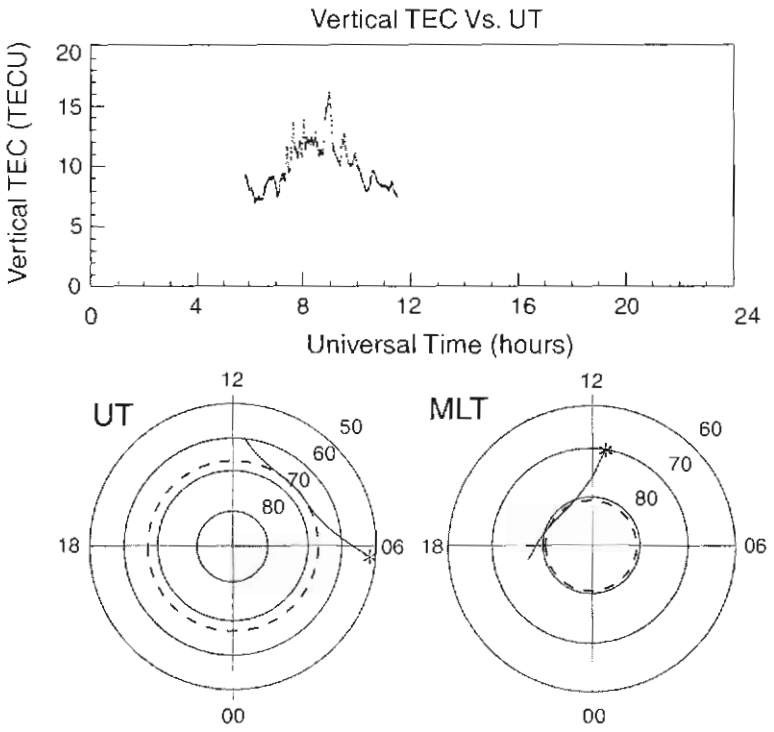


Figure 6. *Vertical TEC and a polar plot of satellite 6 passes in UT/MLT for day 239.*

20.6 Acknowlegements

Thanks to AUSLIG for supplying the GPS TEC data, and a special thanks to the Australian Antarctic Division wintering physicist for the operation of the GISM. This research is being funded by an Antarctic Science Advisory Committee grant with support from the Australian Antarctic Division and the Cooperative Research Center for Satellite Systems (CRCSS). An Australian Postgraduate Award scholarship and the CRCSS support N.M. Shilo.

References

Basu, S., Weber, E.J. and Bishop, G.J. (1990). Plasma structuring in the polar cap. *Journal of Geomagnetism and Geoelectricity*, 42: 763–776.

Beggs, H.M., Essex, E.A. and Rasch, D. (1994). Antarctic polar cap total electron content observations from Casey station. *Journal of Atmospheric and Terrestrial Physics*, 56: 659–666.

Huang, C.S., Andre, D.A. and Sofko, G.J. (1998). Observations of solar wind directly driven auroral electrojets and gravity waves. *Journal of Geophysical Ressearch*, 103: 23347–23356.

Klobuchar, J.A. (1991). Ionospheric effects on GPS. GPS World.

Tate, B.S, Shilo. N.M.S. and Essex, E.A. (1998). An investigation of the ionosphere in the southern high latitude during low sunspot numbers. Acta Geodaetica et Geophysica, Hungarica, 33: 83–90.

Van Dierendonck, A.J., Klobuchar, J. and Quyen Hua (1993). Ionospheric scintillation monitoring using commercial single frequency C/A code receivers. Proceedings of ION GPS-93. Salt Lake City, UT, September 1993.

Van Dierendonck, A.J., Fenton, P. and Klobuchar, J. (1996). Commercial ionospheric scintillation monitoring receiver development and test results. Proceedings of ION 52nd Annual Meeting. Cambridge, MA., June 19-21, 1996.

Weber, E.J. Buchau, J., Moore, J.G ., Sharber, J.R., Livingston, R.C., Winningham, J.D. and Reinisch, B.W. (1984). F layer ionisation patches in the polar cap. *Journal of Geophysical Research*, 89: 1683–1694.

21. USING GPS TO MONITOR IONOSPHERIC IRREGULARITIES IN THE SOUTHERN HIGH-LATITUDE REGION

Yue-Jin Wang, P. Wilkinson and J. Caruana

IPS Radio and Space Services
PO Box 1386
Haymarket NSW 1240 Australia
(email: phil@ips.gov.au)

Abstract

In this paper we discuss a method for using the Global Positioning System (GPS) to monitor ionospheric irregularities in the high-latitude region. Dual-frequency GPS data, obtained from mid-latitude Australian and sub-Antarctic permanent GPS stations, are used to study various ionospheric features, including large scale travelling ionospheric disturbances (TID's) and auroral scintillation. These irregularities are characterised by the presence of rapid changes in the Total Electron Content (TEC). A major magnetic storm period in early November 1997 has been selected to analyse the ionospheric irregularities. The storm, which was associated with an X9-class solar flare, was observed by the Australian ionosonde network. The paper presents the early data analysis results and addresses the technique of using GPS to monitor ionospheric irregularities.

21.1 Introduction

Ionospheric irregularities like Travelling Ionospheric Disturbances (TID's) and ionospheric scintillation have been of increasing technical interest in studying the effects of the ionosphere on communication and navigation systems. TID's are wave-formed structures in the electron concentration of the ionosphere which are caused by movements in the neutral atmosphere. Ionospheric scintillation refers to the amplitude fading or fluctuations in signal phase when radio signals pass through the ionosphere. These two types of ionospheric irregularity are characterised by a rapid change in the total electron content (TEC) (Klobuchar, 1991).

Ionospheric irregularities occur most often in the equatorial region followed by high-latitudes and occur least at mid-latitudes. Equatorial scintillations occur within a belt of ±30° either side of the magnetic equator and are generally seen during local evening hours. In this paper, however, we are mainly concerned with high-latitude irregularities. These irregularities are found in the auroral zone formed in each hemisphere at roughly 64° to 70° geomagnetic latitude. At high-latitudes, the spatial

and temporal extent of irregularities can be much greater than at low-latitudes (Hargreaves, 1992) although the intensity is less.

The Global Positioning System (GPS) radio signals are affected by the ionosphere. By using two frequencies in GPS signals it is possible to measure the ionospheric TEC. In this paper we discuss the method of using GPS to monitor ionospheric irregularities in the high-latitude region. We use the dual-frequency GPS data, obtained from mid-latitude Australian and sub-Antarctic permanent GPS stations, to detect large scale TID's and auroral scintillations during a magnetic storm period in early November 1997. This paper presents the early data analysis results and addresses the technique of using GPS to monitor ionospheric irregularities.

21.2 GPS data

The data used in this study were recorded during 3–8 November 1997 at three Australian Regional GPS Network (ARGN) stations whose geographic locations are illustrated in Figure 1 and listed below.

Tidbinbilla	35.4°S, 148.9°E
Hobart	42.8°S, 147.4°E
Macquarie Island	54.5°S, 158.9°E

ARGN is operated by the Australian Surveying and Land Information Group (AUSLIG). All three stations are equipped with GPS Turborogue dual-frequency receivers recording GPS data with a sampling rate of 30 s.

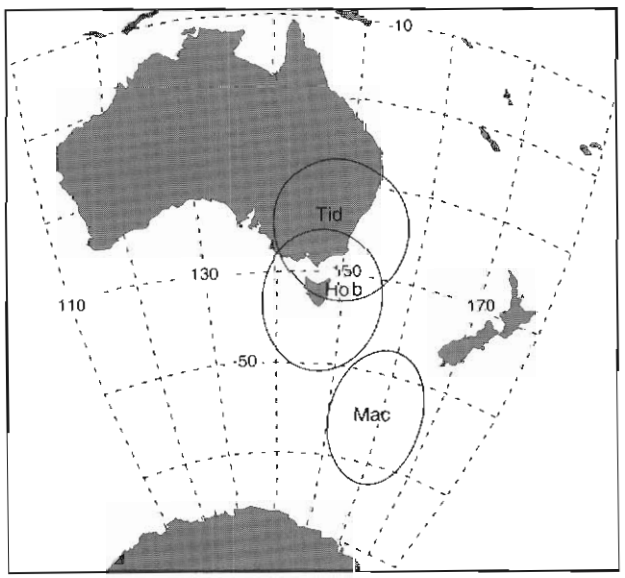


Figure 1. *Map indicating the GPS stations used in the analysis. Circles of range at 20° of elevation are given for each station.*

Both dual-frequency phase and code pseudorange observations have been used to derive the ionospheric TEC. High precision TEC values can be obtained from the carrier phase data, but it contains an initial offset called the initial ambiguity term. To get absolute TEC values, the code pseudorange measurements are used to determine the ambiguity term. By combining GPS carrier phase with code measurements within a satellite pass, the absolute TEC values can be obtained with high precision. The standard procedure to derive TEC values from GPS dual-frequency data can be found from the literature (e.g. Lanyi and Roth, 1988; Wilson and Mannucci, 1993). To eliminate the multipath effects on GPS data, we use only the TEC values for satellites with an elevation angle greater than 20°.

21.3 Magnetic disturbances on 5 November and 7 November 1997

For this study, we selected a solar flare and a period of increasing geomagnetic disturbances, 4–8 November 1997 to investigate ionospheric irregularities at high-latitudes. On 4 November 1997, an X1-flare was observed, followed by an X9-flare on 6 November 1997, which is the largest solar X-ray event since June 1991. A significant geomagnetic disturbance began with a sudden impulse observed by an IPS ground

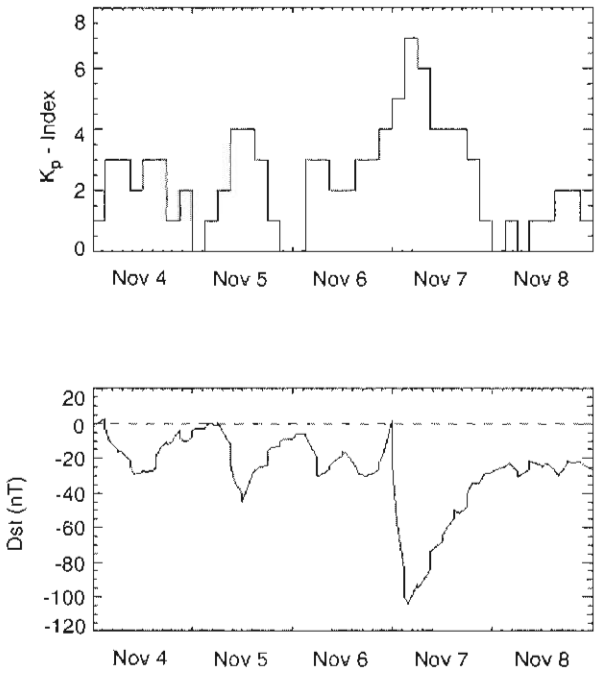


Figure 2. *K_p and D_{st} for 4–8 November 1997.*

magnetometer at 2252 UT on 6 November 1997. Minor to major storm activities were then observed at other sites on 7 November 1997.

In Figure 2, we show data on K_p and the Disturbance Storm Time index (D_{st}) for the period 4–8 November 1997.

Figure 3 shows the vertical TEC values above the three stations during this period. The TEC values are derived by fitting the code pseudorange measurements to the carrier phase. Only data with elevations greater than 65° are used. The TEC values are obtained by averaging the data over every 15 minutes. These TEC values contain instrumental biases in both the satellite and the receiver (Lanyi and Roth, 1988). In our data processing, the receiver biases are corrected by using the receiver bias calibrations provided by Turborouge, while the satellite biases have been obtained from the web site of DLR Fernerkundungs station Neustrelitz in Germany.

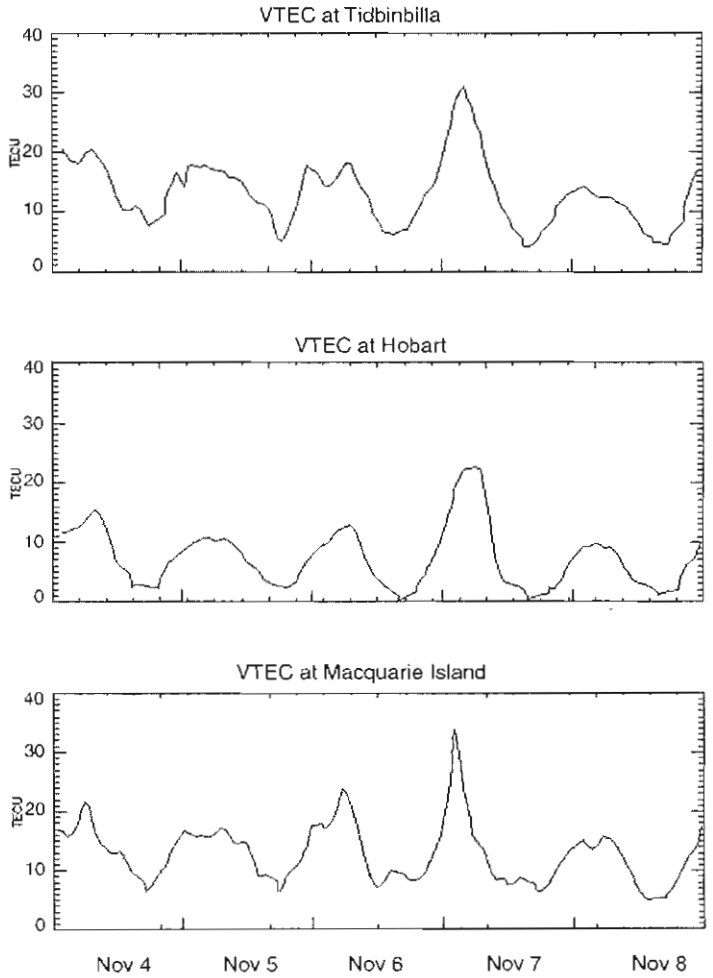


Figure 3. *TEC at Tidbinbilla, Hobart and Macquarie Island for 4–8 November 1997.*

From Figure 3, we can see that the response of the ionosphere to a storm very much depends on the D_{st}. Following the commencement of the storm, the ionospheric response is characterised by an enhancement of ionisation immediately after the storm onset, which is followed by a significant plasma loss on the subsequent storm days, before recovering. This result is consistent with past observations (e.g. Prölss, 1995; Jakowski, 1996). On 7 November 1997, TEC values at Macquarie Island reached a maximum before 0200 UT (1200 local time) while the normal diurnal maximum is around 0400 UT. Figure 3 also shows that Hobart and Tidbinbilla (mid-latitude) have an almost identical trend in TEC enhancement and depression during the magnetically disturbance period.

21.4 Detection of travelling ionospheric disturbances

To study the temporal and spatial variations of ionospheric irregularities, we investigate the rate of change of TEC values covering a large area. It has been suggested that large-scale fluctuations in TEC produce disturbances which propagate through the ionosphere (Hargreaves, 1992).

The TEC variations can be coveniently classified in terms of regular and irregular variations. Regular variations represent the low frequency, long-term trend of TEC, while the irregular variations contain the high frequency, short-term part. In order to isolate short-term TEC changes, we have applied a high-pass filter to the differential carrier phase analysis. The filter has a cutoff of 15 minutes to remove diurnal changes and the variations induced by the changing satellite elevation angle, which are relatively long-term. Short-term changes in ionospheric TEC were computed over a one minute time interval.

The Rate of TEC (ROT) is defined as (Warnant, 1997)

$$ROT^{i}(t_k) = \frac{TEC^{i}(t_k) - TEC^{i}(t_{k-1})}{(t_k - t_{k-1})} \qquad (1)$$

where i denotes the observation to the i^{th} satellite and k is the time epoch. When computing ROT we only use the phase measurements. Since only relative TEC values are involved, the computation of ROT does not require the estimation of the ambiguity term.

Figure 4 displays the phase TEC fluctuations observed at Macquarie Island and Hobart. Large ROT started around 0400 UT at Macquarie Island. About two hours later, corresponding large ROT were observed at Hobart. Possibly, these were TID's passing through the Macquarie Island ionospheric region and then through the Hobart ionospheric region. Hobart is 1544 km to the north of Macquarie Island, suggesting the disturbances travelled with speeds of 200–300 m s^{-1}.

It is noted that these large-scale fluctuations in the TEC due to auroral activity were observed during a daytime post-noon period (0200–0500

local time). Such daytime fluctuations are not seen at equatorial latitudes.

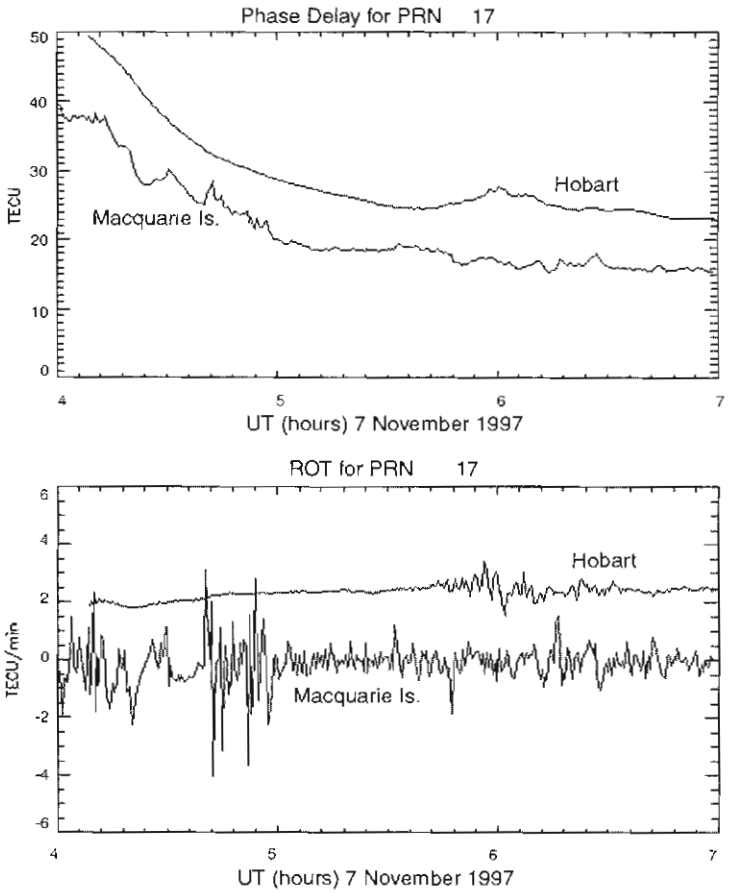


Figure 4. *(top panel) Phase TEC fluctuations at Hobart and Macquarie Island during a magnetic disturbance period. (bottom panel) Rate of TEC at Hobart and Macquarie Island for the same period. The time correlation of the disturbance suggests a TID passing through the ionospheric points above the two stations.*

21.5 TEC fluctuations in the high-latitude region

Large and rapid fluctuations in TEC were observed at Hobart and Macquarie Island during the magnetically disturbed period studied. Figure 5 displays one example of the Macquarie Island TEC fluctuations observed at night on the GPS signals from satellite PRN 1. TEC values for four consecutive days are plotted. The top panel shows the phase fluctuations and the bottom panel displays the ROT. In the panels there

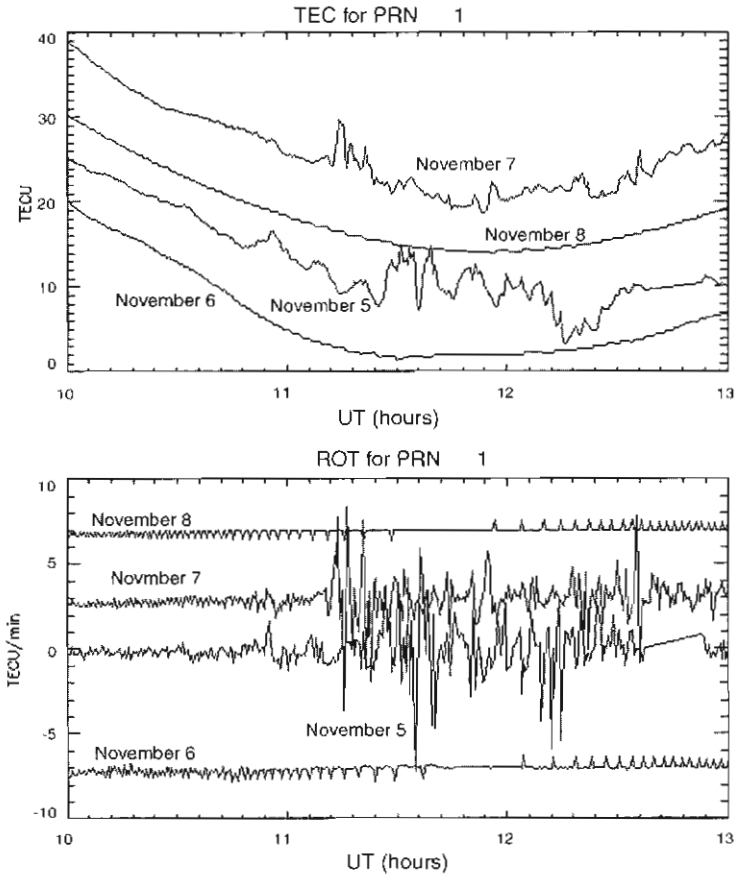


Figure 5. *TEC fluctuations observed at Macquarie Island for satellite PRN 1. Geomagnetic activity was active on 5 November 1997 and a major storm occurred on 7 November 1997. There is an offset applied to each day for clarity.*

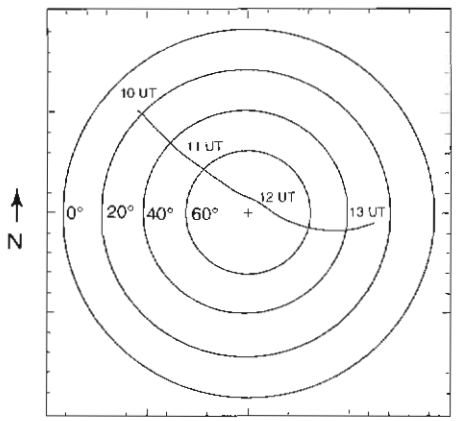


Figure 6. *Azimuth-elevation plots of GPS PRN 1 satellite positions viewed from Macquarie Island on 7 November 1997.*

is an offset applied to each day for clarity. It shows that during the geomagnetically disturbed days, the rate of TEC appears quite irregular and reaches as high as seven TECU m^{-1} in scale. Such a rapid change in TEC may be related to auroral phase scintillations.

Figure 6 shows the satellite position viewed from Macquarie Island station. Since GPS satellites orbit with semi-synchronous periods, their positions repeat from one day to the next (with roughly four minutes shift due to sidereal motion). From the map we can see that the most severe phase TEC fluctuations occurred at a latitude approximately the same as the station (63°S Dip latitude), lasting for more than one hour.

To look at the spatial variation of TEC fluctuations in the high-latitude region, we plotted the Dip latitudes of the sub-ionospheric points where the magnitude of the observed TEC fluctuations are greater than 0.5 TECU m^{-1} against universal time. Figure 7 displays the latitude comparison of TEC fluctuations observed at the two stations. It can be seen that most TEC fluctuations occur at higher Dip latitude (higher than 50°S). The large irregularities follow magnetic activity while few TEC fluctuations were observed during geomagnetically quiet periods, e.g. 8 November 1997.

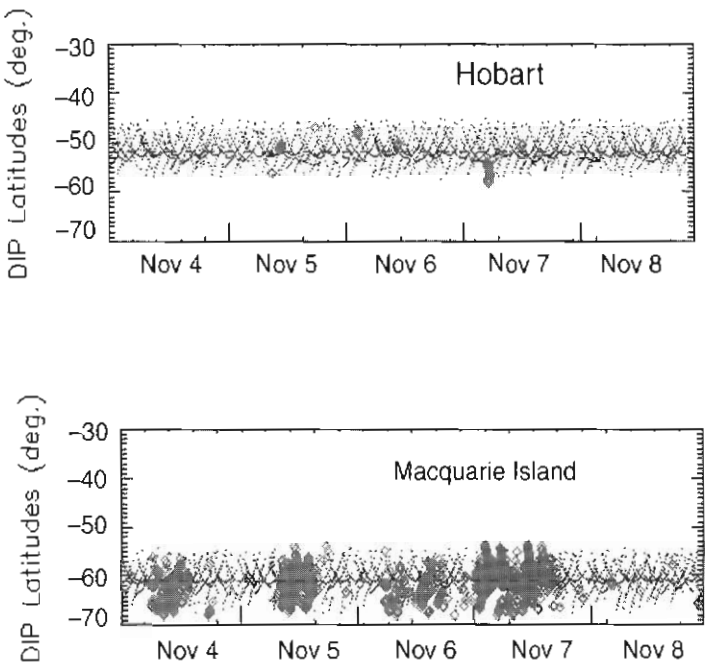


Figure 7. *TEC fluctuations at Hobart and Macquarie Island. Dots show the Dip latitudes of the sub-ionospheric points. The diamond symbol denotes where the magnitude of ROT is greater than 0.5 TECU* m^{-1}.

21.6 Conclusions

We have discussed the method to detect ionospheric irregularities in the high-latitude region using GPS data. This study was undertaken using GPS data gathered during a period of geomagnetic disturbances associated with large solar flares. The results indicate that large TEC fluctuations occur in the high-latitude region during both the daytime and night-time. The extent of ionospheric irregularities is correlated with magnetic activity. By analysing the data from three GPS stations located at different latitudes, ranging from 55°S to 35°S geographically, the large-scale TEC fluctuations due to auroral activity were seen to propagate from Macquarie Island to Hobart at speeds of 200–300 m s^{-1}.

References

Hargreaves, J.K. (1992). The Solar-Terrestrial Environment. Cambridge University Press.

Jakowski, N. (1996). TEC monitoring by using satellite positioning systems. In: Kohl, H., Ruster, R. and Schlegel, K. (Eds.). *Modern Ionospheric Science*. Katlenburg-Lindau, FRG, Pp. 371–390.

Klobuchar, J.A. (1991). Ionospheric effects on GPS. *GPS World*. Pp. 48–51.

Lanyi, G.E. and Roth, T. (1988). A comparison of mapped and measured total electron content using Global Positioning System and beacon satellite observations. *Radio Science*, 23: 483–492.

Prölss, G.W. (1995). Ionospheric F-region storms. In: Vollard (Ed.). *Handbook of Atmospheric Electrodynamics*. CRC Press/Boca Raton. Pp. 195–248.

Warnant, R. (1997). Detection of irregularities in the Total Electron Content using GPS measurements, application to a mid-latitude station. *Beacon Satellite Symposium 1997*. Hungary.

Wilson, B.D. and Mannucci, A.J. (1993). Instrumental biases in ionospheric measurements derived from GPS data. *Proceeding of ION GPS-93*. Salt Lake City, Utah.

22. GPS AND TOPEX MID-LATITUDE TROUGH OBSERVATIONS IN THE SOUTHERN HEMISPHERE AT LOW SUNSPOT NUMBERS

I. Horvath and E.A. Essex

Cooperative Research Centre for Satellite Systems
Department of Physics
La Trobe University
Bundoora Victoria 3083 Australia
(email: e.essex@latrobe.edu.au)

Abstract

For the low sunspot number period of February 1995 to February 1996, the southern hemisphere mid-latitude trough has been studied using the Global Positioning System (GPS) and TOPEX satellite data. A complete procedure was developed for each technique to reduce the raw satellite data to ionospheric Total Electron Content (TEC) values. The GPS data from the ground stations at Tidbinbilla (35.38°S; 148.97°E), Hobart (42.80°S; 147.43°E) and Macquarie Island (54.50°S; 158.94°E) were plotted in various ways to observe variations in trough characteristics that were related to the current magnetic activity, including magnetic storms. The GPS results were supplemented with the TOPEX data, which were also used to establish diurnal and seasonal trends in the Australian region. The feature of initial build-up of ionisation associated with the development of the trough was also investigated. This feature was compared to observations made previously by other researchers utilising different satellite techniques at Macquarie Island. The effects of auroral activity and magnetic storms on GPS recordings and trough occurrence, respectively, were also investigated. The theoretical TEC generated by PIM 1.4, a complex Parameterised Ionospheric Model (Daniell *et al.*, 1995), was compared directly to the GPS TEC.

22.1 Introduction

22.1.1 The mid-latitude trough

The 'main trough' or 'mid-latitude trough' was first reported by Muldrew (1965) and Sharp (1966). It designates an electron depletion zone aligned in the magnetic east-west direction in the nighttime F_2 region of both hemispheres. The trough is located at the magnetic shell parameter L of 4 along the equatorward side of the auroral oval, at an invariant latitude L range of 60°–65° (Taylor, 1973). The trough is associated with two gradients, one equatorward and the other poleward, which define its basic shape. Its average width is around 5°, but the trough can be

up to 10° wide and the factor of electron reduction is between 5 and 10, depending upon the magnetic activity (Schunk *et al.*, 1976). The trough itself terminates the mid-latitude ionosphere and is regarded as a boundary between the mid- and high-latitude regions (Thomas and Dufour, 1965) and as a marker of the plasmapause (Rycroft and Burnell, 1970). The trough is produced by normal F-region processes such as plasma convection, ion production and ion loss. The larger offset between the southern geographic and geomagnetic poles creates hemispherical differences in the Earth's polar environment and therefore there is a marked difference between the northern and southern hemisphere troughs (Fuller-Rowell *et al.*, 1987). Mallis (1989) and Mallis and Essex (1993) investigated the unique characteristics of the southern mid-latitude trough and established diurnal and seasonal trends, at high sunspot numbers. Accordingly, the trough is a persistent feature of the southern ionosphere, appearing in both the daytime and night-time sectors. It is more clearly defined at daytime, regardless of seasons. The narrower and deeper night-time trough lasts for longer periods of time, even after midnight. The night-time trough was observed more frequently and better developed in the seasons of summer and vernal equinox, than in any other season of the year, making up the highest frequency of occurrence and greatest prominence. These features became increasingly moderate through the autumnal equinox toward winter.

22.1.2 Aim

The principal aim of this project was to study the main characteristics of the southern hemisphere mid-latitude trough and the structure of the ionisation at the time of trough occurrence at low sunspot numbers, by utilising two independent satellite techniques, namely GPS and TOPEX. The total electron content values obtained from raw satellite data were used to identify the location, the time and rate of development and the spatial configuration of the trough. The information collected was used to establish diurnal, seasonal and magnetic activity related trends.

22.1.3 Study area

This study makes use of GPS measurements recorded from a small number of receiver stations forming a longitudinal chain across the mid-latitude Australian continent to the high-latitude sub-Antarctic region, in order to derive land-based total electron content values (see Figure 1(a)). The receiver stations are situated at the locations of Tidbinbilla, Hobart and Macquarie Island. For the comparison of results of the different techniques, long and continuous TOPEX passes were selected from the Australian sector of the Pacific Ocean, situated close to the east coast section of the Australian GPS receiver network (see Figure 1(b)).

22.1.4 Database

The GPS database was established by collecting RINEX observational and TEC data files through the Internet for up to seven days, centred on the Priority Regular World Day of each month, for a 13-month period extending between February 1995 and February 1996 inclusive. The GPS database was augmented with the extensive collection of TOPEX data,

stored on CD-ROM's and each containing two complete repeating cycles, covering the same period of time.

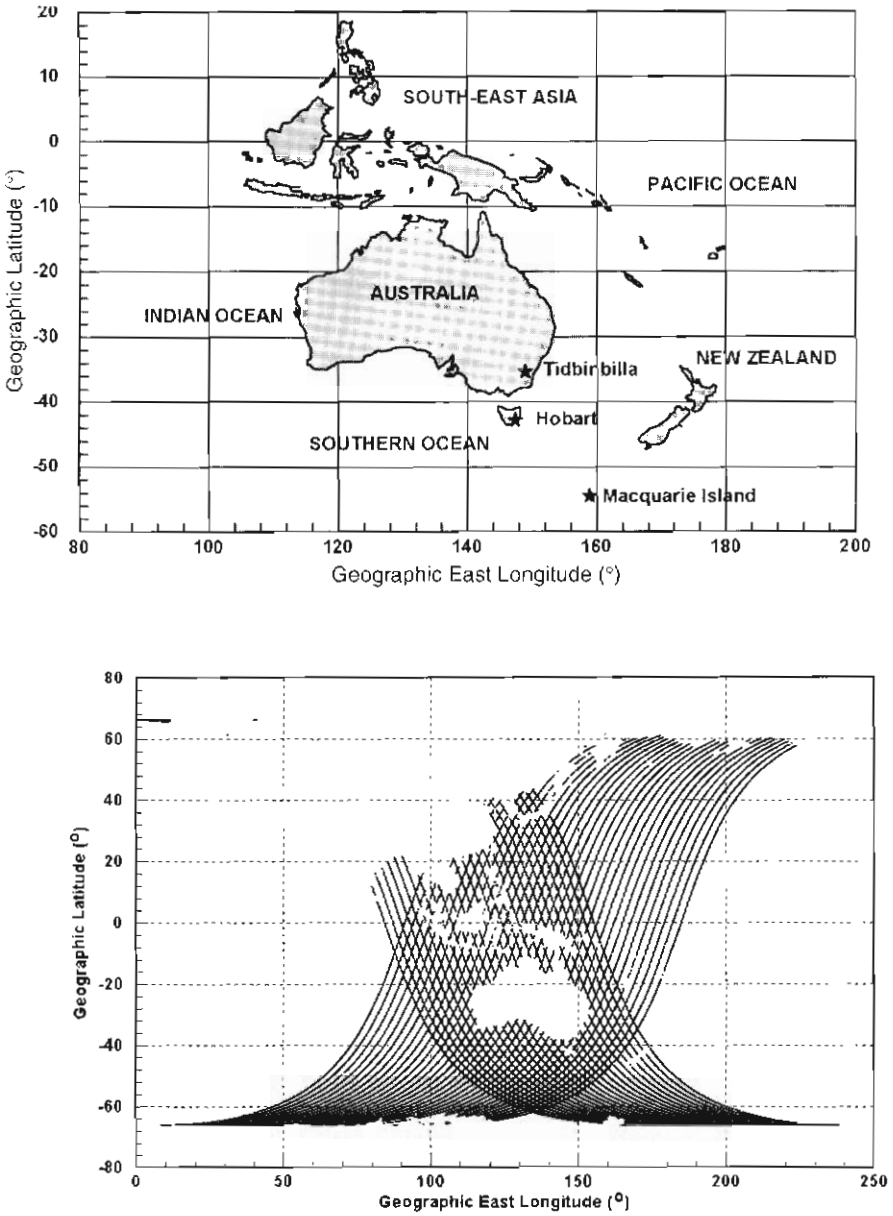


Figure 1. *(a) The map of the GPS study area illustrates the location of the dual-frequency receiver sites in the grid of geographic co-ordinates. (b) The ground tracks of the TOPEX/Poseidon satellite define the TOPEX study area over the oceans in the Australian region.*

22.2 Experimental and theoretical considerations

The dual-frequency GPS users take advantage of two of the ionospheric effects occurring on modulated L-band radio signals propagating to Earth. One is the absolute differential time delay ($\delta\Delta T$) measured between the two carriers (in nanoseconds), which allows one to compute the value of differential group path delay ($\delta\Delta P'$ in meters). The other is the relative differential carrier phase advance ($\delta\Delta f$), recorded with respect to the lower frequency (in cycles). The combination of $\delta\Delta P'$ and $\delta\Delta f$ gives the slant GPS TEC (see Equation 1()) after the magnitude of the baseline, i.e. the offset between the two data curves (see Figure 2), was established (Klobuchar, 1996).

$$\text{slant GPS TEC} = \int_{O}^{S} N\,ds = \frac{2cf_2}{K}\left(\frac{f_1^2}{f_1^2 - f_2^2}\right)\delta\Delta\phi + \text{baseline} \qquad (1)$$

where: $c = 2.997 \times 10^8$ m s^{-1}, $K = 80.62$ m^3 s^{-2}, $f_1 = 1.57542$ GHz and $f_2 = 1.22760$ GHz.

A vertical content can be computed if the orbital elements of the satellite are known. Since the orbital height of the satellite (20 183 km) is in the range of the plasmasphere, the vertical content obtained is the sum of ionospheric (ITEC) and protonospheric (PTEC) components:

$$\text{vertical GPS TEC} = \text{slant GPS TEC} * \overline{\cos\chi} = \text{ITEC} + \text{PTEC} \qquad (2)$$

where: $\overline{\cos\chi}$ = mean value of $\cos\chi$ at the median height of the ionosphere along the integration path.

χ = zenith angle of the ray.

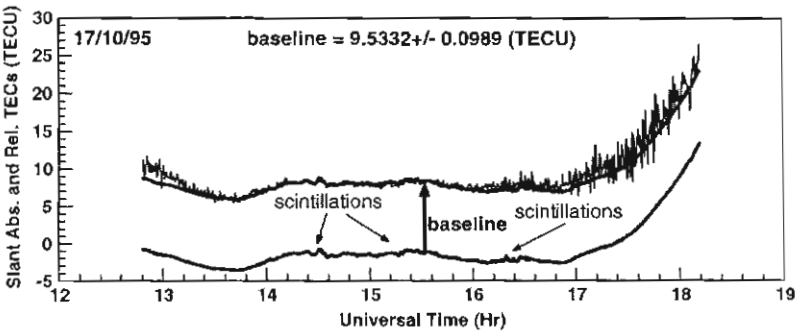


Figure 2. *The Macquarie Island GPS plot shows that the group path and carrier phase measurements give TEC values, which represent a less accurate absolute scale (top curve) and a more accurate relative scale (bottom curve), respectively. The concept of baseline is also shown. The scintillations of the carrier phase signal, as a response to auroral disturbances, are also indicated.*

The ionospheric height correction ($\Delta R_{ionosphere}$) (see Equation (3)), made by the on-board radar of the TOPEX/Poseidon satellite taking vertical sea height measurements at Ku-band and C-band frequencies, gives unambiguous TEC measurements (Imel, 1994).

$$\Delta R_{ionosphere} = -\frac{1}{2}\frac{K}{f^2}\int_0^{h_s} N dh \qquad (3)$$

where: f = 13.65 GHz.

Since the altimeter operates in the nadir direction and the orbital height (1336 km) of the satellite is in the topside F_2 region, a vertical ionospheric TEC can be obtained directly:

$$\text{TOPEX TEC} = \frac{\Delta R_{ionosphere} f^2}{-403.1} = \text{ITEC} \qquad (4)$$

where: TOPEX TEC is in TECU, $\Delta R_{ionosphere}$ is in mm and f is in GHz.

22.3 Experimental results and discussion

As the trough is a structure not only in space but time, as well, its main characteristics will be discussed and illustrated by vernal equinox GPS data, plotted in time and geographic co-ordinates. The GPS findings will be compared with the TOPEX results. As Figure 3(a) shows, on 20 September 1995 the trough was detected from Macquarie Island at around 1500 LT and remained visible for several hours. The main features of the trough such as the initial build-up of ionisation and the equatorward wall are shown in Figures 3(a) and (b). Here the satellite PRN 7 detected the equatorward wall twice, since the satellite turned back to lower latitudes after reaching its maximum position at the station. The equatorward and poleward walls and the trough minimum are clearly visible in Figures 3(a) and (c). Here the satellite PRN 6 travelled straight through the trough minimum and later turned south of Macquarie Island. Therefore it detected the equatorward and poleward walls, and a small section of the auroral oval, indicated by the fine-scale TEC fluctuations. The signatures of auroral activity appear as scintillations on the carrier phase data (see Figure 2) and indicate that the satellite beam travelled through the auroral oval. Observations from the low sunspot number GPS data revealed that the initial build-up of ionisation is a distinctive and regular feature developing on the equatorward wall of the trough. It can be defined as a sudden increase in TEC from normal to anomalously high values that peak and drop back at the edge of the electron depletion. According to Foster (1993), fresh co-rotating ionospheric plasma along the equatorward edge of the two-cell convection pattern produces such increases. The TOPEX trough observations indicated also the initial build-up of ionisation and therefore confirmed the GPS observations. Two of those TOPEX observations are shown in Figures 4(a) and (b). The TOPEX TEC is plotted against the geographic latitude and shows a vernal equinox trough on 28 September 1995 and 4 October 1995, at around 15.7 LT.

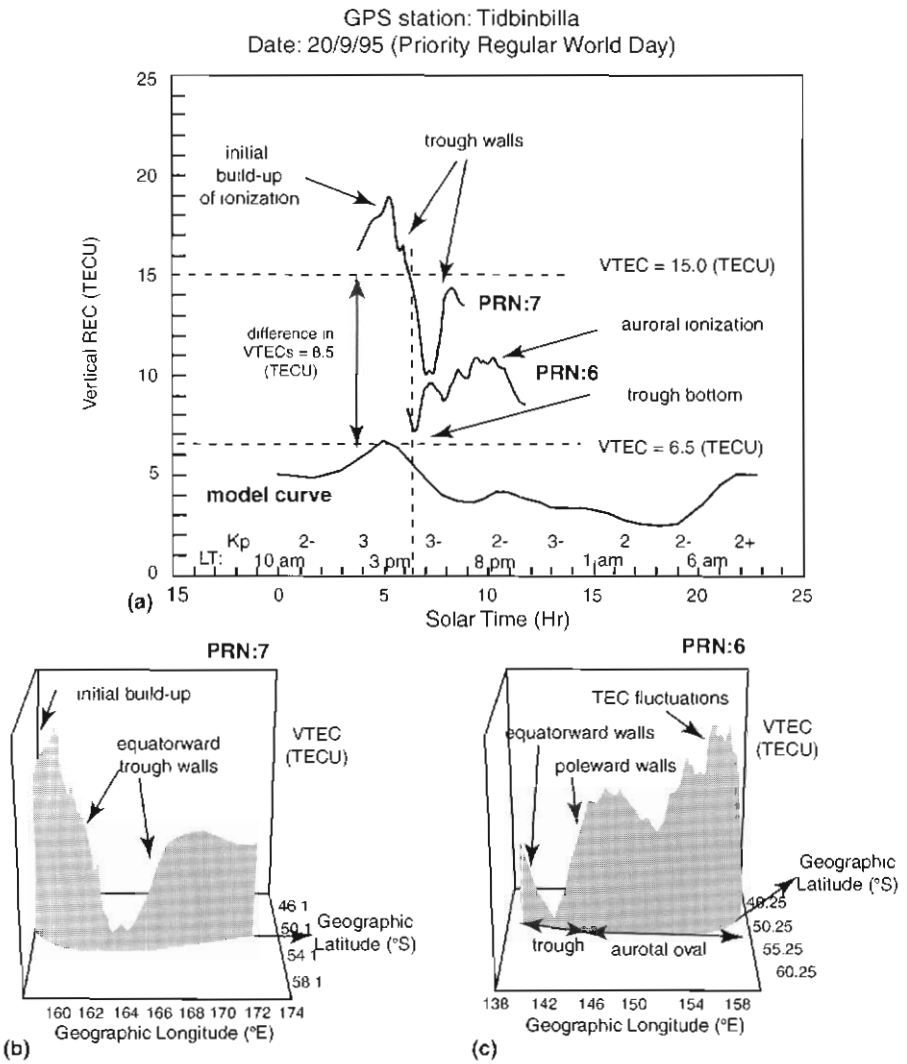


Figure 3. *(a) The 2-dimensional daily and (b and c) 3-dimensional spatial GPS TEC plots depict the major characteristics of a daytime vernal equinox trough. The direction of the satellite pass is also shown (–>). (a) A PIM generated model TEC curve is displayed for comparing theoretical and experimental results. (a, b and c) The satellites are designated by their PRN (Pseudo Random Noise) numbers.*

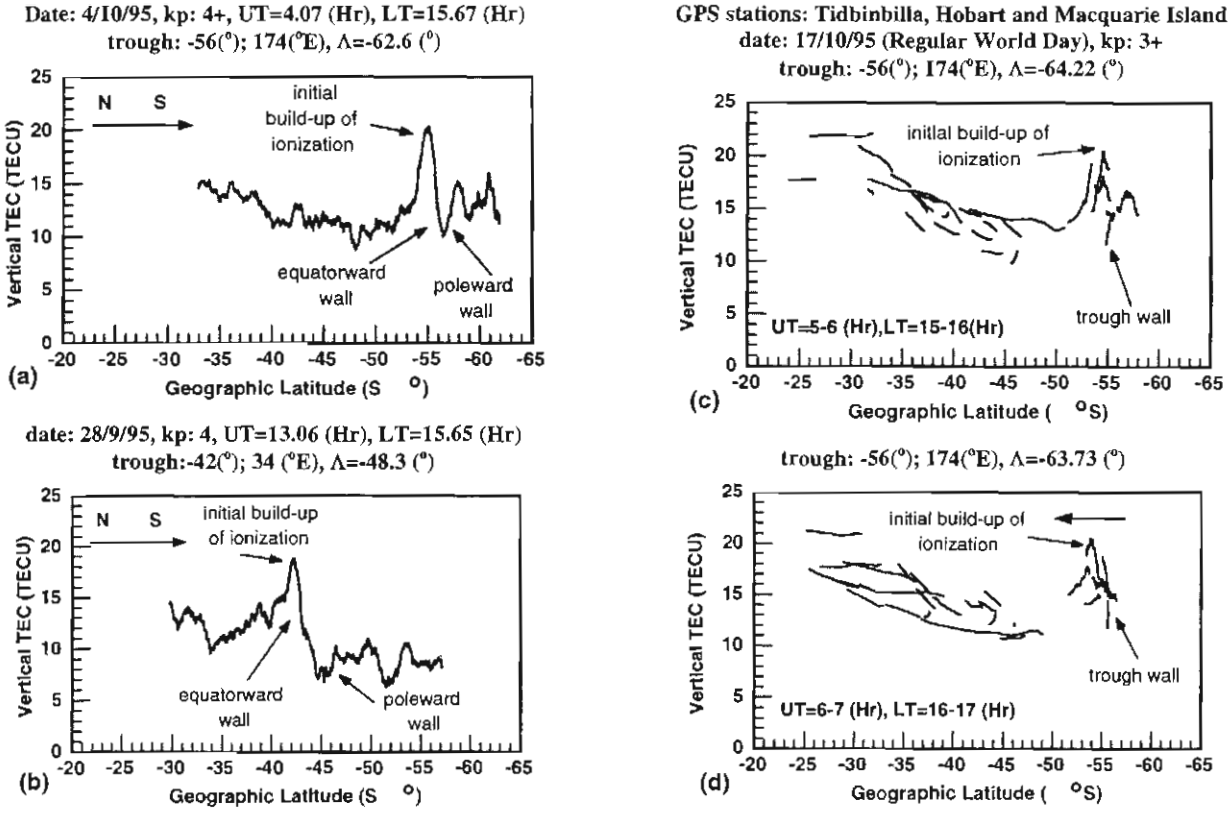


Figure 4. *(a and b) The TOPEX maps depict a daytime vernal equinox trough. The direction of the satellite pass is also shown (→). (c and d) The hourly GPS maps show the diurnal latitudinal movement of a daytime vernal equinox trough.*

For the period investigated, the diurnal variation of the trough shows a higher number of daytime and a smaller number of night-time trough occurrence. The trough develops into a more well-defined formation in the daytime sector. The location of the trough was studied with hourly latitudinal maps. The GPS plots on 17 October 1995 at 1500–1600 LT and 1600–1700 LT show that the peak of the initial build-up moves towards the equator at 160°E longitude, after the trough appeared in the daytime sector (see Figures 4(c) and (d)). These latitudinal maps also show the gradual and linearly increasing electron content toward the equator, from ordinary low values at higher mid-latitudes. The peak TEC related to the initial build-up of ionisation is almost as high as the maximum values observed at the low-latitude region.

It is interesting to compare trough observations obtained by different techniques operating at the same recording station, during different periods of the solar cycle. The TEC data from the Faraday technique at Macquarie Island (Lambert and Cohen, 1986; Lambert and Essex, 1987), in a low sunspot number period, show the trough for vernal equinox in 1984 at 0600 UT with the initial build-up of ionisation on the equatorward side (see Figure 5(a)). Contrary to this, the differential phase technique, employed in a high sunspot number period (Mallis, 1989; Mallis and Essex, 1993), detected the trough without this distinctive characteristic on vernal equinox 1988 at around 4.7 UT (see Figure 5(b)).

The most significant seasonal trend observed is, that the daytime trough develops best during equinoctial times, less in summer and least in winter. During the summer season the night-time trough develops better, which is narrower and deeper than its shallow and wide daytime counterpart.

The trough is a structure in space and time, and has a substantial movement during the course of the year. This is called annual latitudinal movement and was studied with TOPEX passes (see Figure 6) of similar equator crossings related to different seasons. Figure 6(a) shows a summer trough on 15 February 1995 at 1459 LT with the equatorial anomaly centered on the geomagnetic equator. Figure 6(b) depicts a winter trough on the 20 July 1995 at 1411 LT and shows the equatorial anomaly in a not so well developed form at 0918 LT. During those moderately disturbed magnetic periods, the daytime trough was observed at higher latitudes ($\approx 60°S$, $\Lambda = -73.8°$) in summer than in winter ($\approx 52°S$, $\Lambda = -62.4°$). In agreement with the results of Mallis (1989) and Mallis and Essex (1993), this indicates a movement away from the equator and then back towards lower latitudes, during the course of the year. Figure 6 also illustrates that in winter, the daytime trough is shallower than in summer, because of the less ionisation available.

The effects of magnetic disturbance on trough occurrence and GPS recordings were studied with GPS TEC data. The magnetic storm studied, identified from geomagnetic data, commenced on 18 October 1995 at 1400 UT. On that day the minimum D_{st} index was -122 nT and the maximum K_p was 7^-. During the storm, the trough was detected from

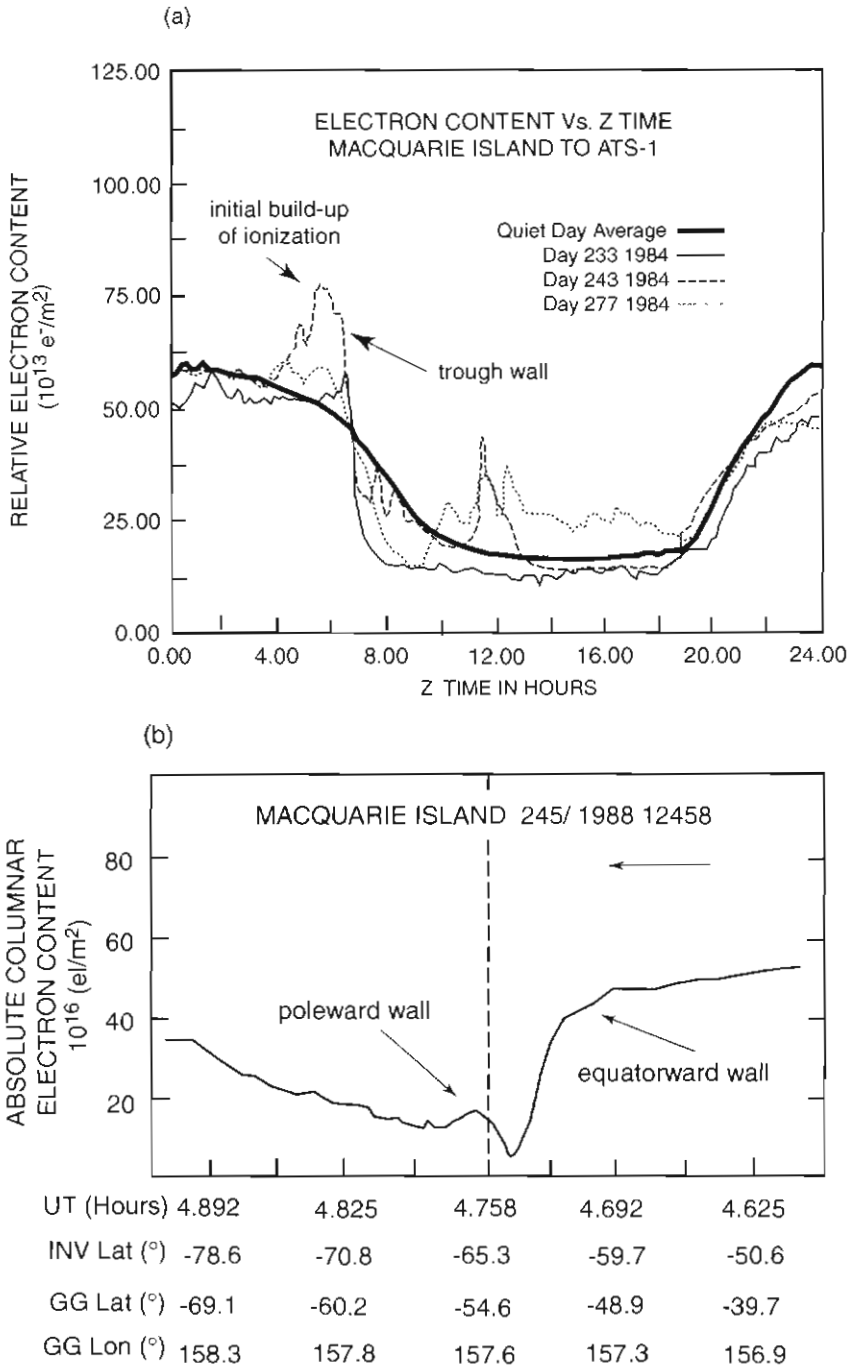


Figure 5. *The plots of TEC from (a) the Faraday technique (modified after Lambert and Essex, 1987) and (b) the differential phase technique (modified after Mallis, 1989) show the mid-latitude trough at Macquarie Island in a low and a high sunspot number period, respectively. (b) The direction of the satellite pass is also shown (←).*

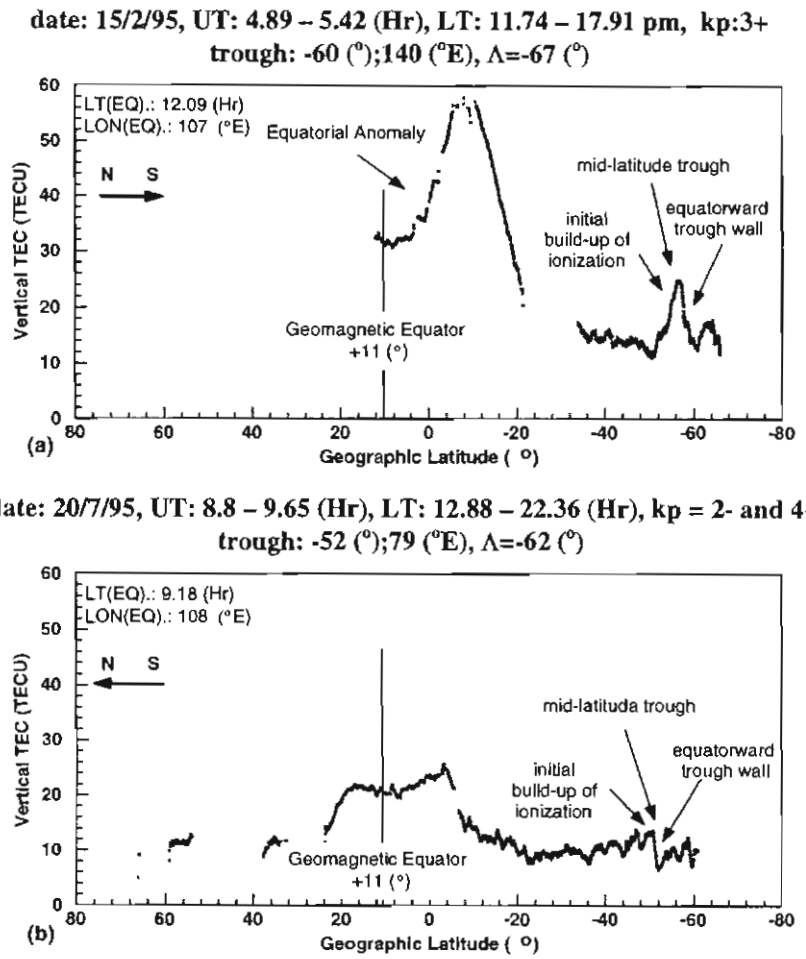


Figure 6. *The TOPEX passes depict a daytime southern mid-latitude trough in the seasons of (a) summer with the southern half of the equatorial anomaly, and (b) winter, and indicate the annual movement of the trough. (a and b) The equator crossing local time (LT(EQ)), the equator crossing longitude (LON(EQ)) and the direction of the passes ((a) → and (b) ← are also shown.*

Hobart, situated at lower geographic latitude than Macquarie Island, indicating that the trough follows the movement of the expanding auroral oval. At that time fine- and larger-scale TEC fluctuations and large-scale wave like formations were also detected. Figure 7 shows these ionospheric formations on a GPS TEC latitude plot constructed for a one-hour period (LT = 1800–1900) on 19 October 1995. Since the different segments of the trough are detected by different satellites, its complete image is highlighted, in order to obtain a clearer presentation. Opposite

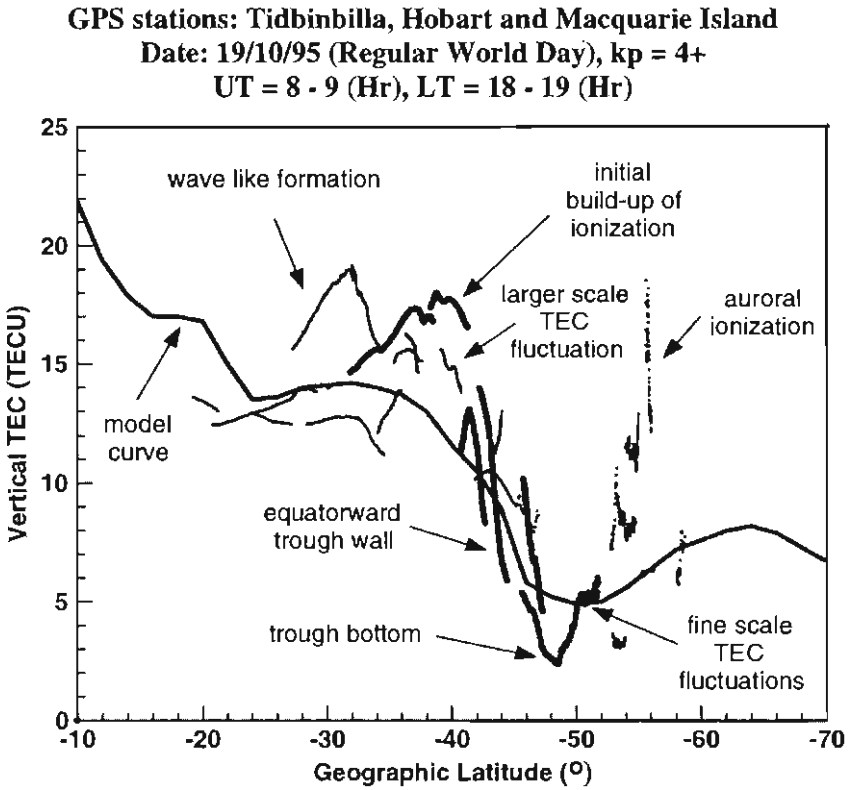


Figure 7. The latitudinal GPS plot depicts a vernal equinox daytime trough detected from Hobart during the 19 October 1995 magnetic storm. The PIM generated model TEC is also shown.

to storm times, during less disturbed magnetic periods, the auroral oval is situated at higher latitudes and the trough is observed from the location of Macquarie Island in the Southern Ocean.

The theoretical TEC values generated by PIM were directly compared to the GPS results, for different magnetic conditions. In Figure 3(a) during a moderate magnetic disturbance, the model values are significantly lower than the experimental values and the phenomenon of mid-latitude trough is not indicated at all. Figure 7 shows that the agreement is better for higher K_p, but PIM is still not able to model the trough.

22.4 Conclusion

The overall results obtained in this project indicate that the combination of dual-frequency GPS and TOPEX techniques provide a comprehensive way to study the ionosphere. Approximately one year of GPS data was analysed to study the major characteristics of the mid-latitude trough of the southern hemisphere, during low sunspot numbers. The GPS findings were confirmed with TOPEX observations, which provided additional information on the structure of the ionisation at the time of the trough occurrence. The initial build-up of ionization was found to be a unique feature that was also detected by the Faraday method at low sunspot numbers, but was not detected by the differential-phase technique during high sunspot number periods. The high number of daytime observations indicates that the trough develops better in the daytime sector. The trough was found to be better defined in the equinoctial seasons, while the degree of development and the frequency of occurrence were observed to be decreasing through the summer towards winter. The daily and annual movements of the trough were also defined. The established diurnal and seasonal trends are in good agreement with the findings of Mallis (1989) and Mallis and Essex (1993). The magnetic activity related trends were also investigated and the findings supported the well-known mechanism of the polar environment. The experimental data were compared with the PIM generated theoretical TEC. For the low sunspot number period, the theoretical TEC is generally below the measured values and PIM is not able to model the structure of the ionisation at the time of trough occurrence.

22.5 Acknowledgements

This research is supported by an Australian Telecommunication and Electronics Research Board (ATERB) grant and an Antarctic Science Advisory Committee (ASAC) grant. I. Horvath is supported by a La Trobe University Postgraduate Award (LUPA) and the Cooperative Research Centre for Satellite Systems (CRCSS). Special thanks are extended to IPS Radio and Space Services, AUSLIG and JPL for data. The authors thank P. Doherty from Boston College USA, J. Klobuchar from ISI USA and Dr A. Breed from the Australian Antarctic Division for assistance and advice with this project.

References

Daniell, R.E. Jr., Brown, L.D., Anderson, D.N., Fox, M.W., Doherty, P.H., Decker, D.T., Sojka, J.J. and Schunk, R.W. (1995). Parameterized ionospheric model: a global ionospheric parameterization based on first principles models. *Radio Science*, 30: 1499–1510.

Foster, J.C. (1993). Storm related transport at middle and high latitudes. *Journal of Geophysical Research*, 98: 1675–1689.

Fuller-Rowell, T.J., Rees, D., Quegan, S., Moffett, R.J. and Baily, G.J. (1987). Interactions between neutral thermospheric composition and the polar ionosphere using a coupled ionosphere-thermosphere model. *Journal of Geophysical Research*, 92: 7744–7748.

Imel, D.A. (1994). Evaluation of the TOPEX/Poseidon dual frequency ionospheric correction. *Journal of Geophysical Research*, 99: 24895–24906.

Klobuchar, J.A. (1996). Ionospheric effects on GPS. In: Parkinson, B.W. and J.J. Spilker, J.J. (Eds.). *Global Positioning System: theory and applications*. American Institute of Aeronautics and Astronautics Inc., Washington. Pp. 485–515.

Lambert, M. and Cohen, E.A. (1986). Monitoring ionospheric irregularities in the southern auroral region by means of a satellite beacon. *Radio Science*, 21: 347–350.

Lambert, M. and Essex, E.A. (1987). Studies of ionospheric irregularities around L=4 in the southern hemisphere using satellite beacons. *ANARE Research Notes 48*. Australian Antarctic Division. Pp. 198–205.

Mallis, M. (1989). *An investigation of the southern-hemisphere ionospheric main trough using total electron content measurements*. Ph.D. Thesis, La Trobe University.

Mallis, M. and Essex, E.A. (1993). Diurnal and seasonal variability of the southern-hemisphere main ionospheric trough from differential-phase measurements. *Journal of Atmospheric and Terrestrial Physics*, 55: 1021–1037.

Muldrew, D.B. (1965). F-Layer ionization trough deduced from Alouette data. *Journal of Geophysical Research*, 70: 2635–2650.

Rycroft, M.J. and Burnell, S.J. (1970). Statistical analysis of movements of the ionospheric trough and the plasmapause. *Journal of Geophysical Research*, 75: 5600–5604.

Sharp, G.W. (1966). Midlatitude trough in the night ionosphere. *Journal of Geophysical Research*, 71: 1345–1356.

Schunk, R.W., Banks, P.M. and Raitt, W.J. (1976). Effects of electric fields and other processes upon the nighttime high-latitude F layer. *Journal of Geophysical Research*, 81: 3271–3282.

Taylor, H.A. Jr. (1973). The light ion trough. *Planetary Space Science*, 20: 1593–1605.

Thomas, J.O. and Dufour, S.W. (1965). Electron density in the whistler medium. *Nature*, 206: 567–571.

23. MONITORING THE IONOSPHERE/PLASMASPHERE WITH LOW EARTH ORBIT SATELLITES: THE AUSTRALIAN MICROSATELLITE FEDSAT

E.A. Essex, P.A. Webb, I. Horvath, C. McKinnon, N.M. Shilo and B.S. Tate

Cooperative Research Centre for Satellite Systems
Department of Physics
La Trobe University
Bundoora Victoria 3083 Australia
(email: e.essex@latrobe.edu.au)

Abstract

Recently, with the escalating cost of large satellite missions, attention has turned to smaller satellites. Their advantages of low overall cost in construction and launch, and short time span between conception and launch has given a new impetus to the further study of the geosphere. By using a combination of space-based and ground-based receivers, it is possible to undertake new and exciting experiments directed towards furthering our knowledge of the ionosphere. Combinations of high Earth orbit satellites, such as the Global Positioning System (GPS), and Low Earth Orbit (LEO) microsatellites are providing the capability for satellite-to-satellite occultation experiments to reconstruct the vertical profile of the ionosphere. The topside ionospheric and plasmaspheric ionisation content may also be explored with satellite-to-satellite experiments. This paper describes some of the experiments proposed for future microsatellites, such as those planned for the Australian satellite FedSat.

23.1 Introduction

The ionised part of the atmosphere, the ionosphere, causes distorting effects or errors in satellite communications, navigation and altimetry. In order to understand these effects and improve our knowledge of the ionosphere, the various regions of the ionosphere have been monitored to different degrees on a long-term basis. Even instruments such as incoherent scatter radars, while providing profiles of both the topside and bottom side ionosphere, do not operate on a continuous basis and provide poor resolution for the lower ionosphere. The ground-based ionosondes provide non-homogeneous coverage of the bottom side ionosphere, as they are limited to observations from land. This is particularly true in the southern hemisphere where the oceans cover most of the hemisphere. The topside ionosphere and the region above, the plasmasphere are not easily monitored on a long-term basis. Only a few instruments such as incoherent scatter radars, topside sounders and *in situ* satellite-based diagnostics are able to provide details on the structure, composition and dynamics of the topside ionosphere. The

properties of the plasmasphere are even less well known, and techniques to explore it are very limited. The advent of the Global Positioning System (GPS) with 24 satellites in 12 hour orbits at 20 000 kilometres, provides the opportunity to monitor on a global basis the variation of the ionisation content of the ionosphere and plasmasphere. For the GPS navigation and timing systems, the ionosphere produces the largest source of errors. These errors can be measured and utilised to improve models of the global distribution of ionisation, and hence be used to improve error corrections in future satellite applications. The GPS receiver network on the ground, together with GPS receivers in orbit, provide the unique opportunity to perform this task. One of the payloads on the Australian scientific microsatellite, FedSat, planned for launch in November 2001, is a GPS receiver. The launch coincides with the expected peak of the sunspot cycle, and hence at this time the disruptions to satellite reception and error corrections due to space weather are predicted to be at maximum levels. The recent deployment of space-based GPS receivers has demonstrated the feasibility of using satellite-to-satellite experiments for monitoring the ionosphere. In this paper we discuss some of the satellite-to-satellite experiments proposed for the Australian satellite FedSat.

23.2 FedSat

FedSat, the first Australian satellite to be launched in over 30 years, will be a microsatellite of around 58 kg launch mass. It is planned to launch into a Sun synchronous polar orbit at a height of 800 km. FedSat's mission is basically a science and engineering one. The planned scientific payloads include a fluxgate magnetometer for monitoring the Earth's magnetic field and solar terrestrial interactions as well as a dual frequency GPS receiver for ionospheric and atmospheric research. The aims of the program are to conduct basic research on the structure and dynamics of the ionosphere, plasmasphere and magnetosphere. The results will be applied to the forecasting of space weather. The space-qualified dual-frequency GPS receiver and patch antenna, supplied by NASA, are being built by Spectum Astro of Arizonia, USA. Using the two coherently connected frequencies f_1 = 1575.42 MHz and f_2 = 1227.60 MHz, high precision measurements will be made of the group and phase delay between the receiver on board FedSat and the transmitters on the GPS satellites visible to FedSat. Two types of measurements are planned. In the occultation mode (Hajj *et al.*, 1994), the ray path from the GPS satellites will pass horizontally through the ionosphere and hence provide scans through the ionosphere as FedSat rises and sets with respect to a GPS satellite (Hajj and Romans, 1998). The second type of measurement is for the overhead mode where the ray paths from the GPS satellites in a vertical cone above FedSat pass through the plasmasphere. Figure 1 illustrates the two modes. As the FedSat orbit is planned to be at 800 km, most of the ionisation measured by this technique is located in the plasmasphere and topside ionosphere. The following sections detail some of the proposed experiments to be carried out using the on board GPS receiver.

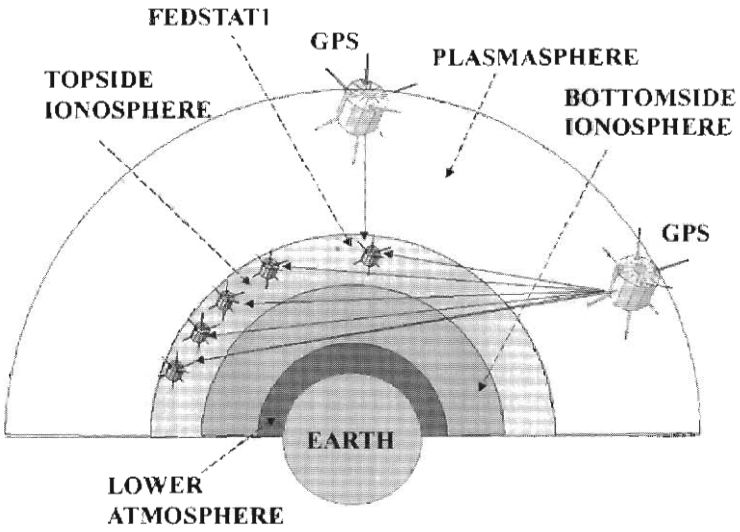


Figure 1. *Model (not to scale) of the ray paths from the GPS satellites to the receiver on board FedSat. The occultation mode and the overhead mode are illustrated.*

23.3 Ionospheric measurements

The application of limb sounding in planetary occultation experiments has been used repeatedly by NASA to study the atmospheres of planets in our solar system (Tyler, 1987). The application of the technique to studies of the Earth's atmosphere had to wait until a suitable radio source, namely GPS, became available. Observing the GPS constellation of satellites in space from a Low Earth Orbit (LEO) satellite such as FedSat, orbiting at 800 km provides a powerful means of imaging the ionosphere and the inner magnetosphere, the plasmasphere. The provision of occulting geometry from a space-based receiver enables effectively a horizontal scan through the ionosphere to be obtained (Hajj *et al.*, 1994). The lack of horizontal scans from ground-based observations of GPS satellites has limited the current development of computerised ionospheric tomography. Because of the geometry, there is a lack of horizontal Total Electron Content (TEC) information that causes a lack of information along the vertical direction in the reconstructed tomographic image. To overcome this problem most researchers use some form of *a priori* information (Villani and Essex, 1996). A review of ionospheric tomography algorithms may be found in Raymund (1994).

In the planned experiment for FedSat, this problem will be overcome by using a combination of ground-based as well as space-based GPS occultation data which will provide both vertical and horizontal information for the application of the tomographic reconstruction (Horvath and Essex, 2000). Although in one day there will be many hundreds of occultations of FedSat with GPS satellites, only those close to the track of the FedSat satellite will be used. Off track measurements are difficult to interpret as they move through a large range of latitude and longitude. As FedSat will be in a Sun synchronous polar orbit, it will be possible to study not only the low and mid-latitude ionosphere, but also the higher-latitude ionosphere especially over the vast expanses of the oceans in our region. The principal advantages of the tomographic reconstruction techniques lie in the large geographic coverage and the cost effectiveness. Incoherent scatter radars are only able to provide the information from a smaller geographical area at the cost of millions of dollars to build and operate. Recent interest in the development of space-based tomography has seen the launch of the first GPSMET LEO, in the proof of concept for Earth radio occultation experiments. Various researchers have recently performed ionospheric reconstruction from the GPS occultation data (Ruis *et al.*, 1997; Hajj and Romans, 1998).

23.4 Plasmasphere measurements

A further area of interest is the electron content of, as well as the irregularities in, the plasmasphere between FedSat and the GPS satellites orbiting above FedSat. With a GPS antenna located on the top of the FedSat satellite, it will be possible to undertake TEC measurements vertically above FedSat to the GPS satellites for coincidence occurrences in the orbits. For antenna reception within a cone of five degrees above FedSat these coincidences would enable information to be obtained on the topside ionosphere and the plasmasphere. Figure 2 shows model calculations of the TEC for three cases: below 400 km, between 400 and 800 km, and above 800 km. The model plasmasphere used in these calculations is based on diffuse equilibrium, and in combination with the International Reference Ionosphere (IRI), forms a global ionospheric-plasmasphere model (Webb and Essex, 1997). At low-latitudes at sunspot maximum, there is clearly significant ionisation above 400 km. For the high-latitudes in the Australian region of the southern hemisphere, where the magnetic pole is offset from the geographic pole toward Australia, the plasmapause is located at lower geographic latitudes. Hence for coincidences at higher latitudes, it will be possible to estimate the location of the plasmapause. This is because, at these locations, the ray paths from the GPS satellites to the FedSat satellite will not intersect the plasmasphere, and there would normally be only a negligible amount of ionisation above 800 km, so that the TEC should approach zero (Klobuchar *et al.*, 1994). Recent analysis of ground-based GPS TEC data has shown the existence of a large enhancement of ionisation equatorward of the mid-latitude trough in the plasmapause region. This increase in ionisation may be related to heating of the topside ionosphere by turbulent dissipation of ring current energy at the

plasmapause (Horvath and Essex, 2001; Titheridge, 1976). FedSat may also be used to investigate this ionisation buildup, especially over the southern hemisphere where the ground-based measurement are sparse.

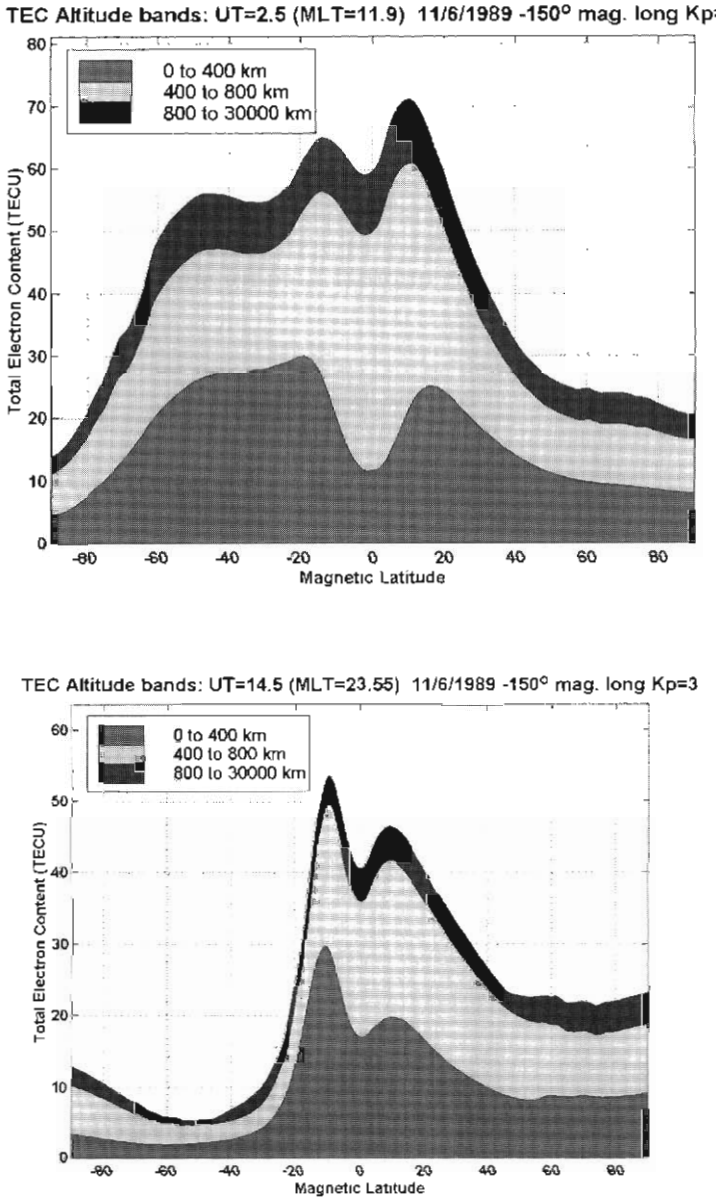


Figure 2. *Electron content contributions from 0 to 400 km, 400 km to 800 km and from above 800 km to the TEC for a GPS satellite orbit for a longitude slice at 1500 magnetic longitude at midday and midnight. The date corresponds to a southern hemisphere winter sunspot maximum period from the last sunspot cycle.*

23.5 Irregularities

Irregularities in the plasma between FedSat and the GPS satellites should also be detectable as scintillation activity in the phase path data. These irregularities may be located in either the ionosphere or the plasmasphere. Jacobson *et al.* (1996) has used sensitive ground-based techniques to investigate field-aligned irregularities in the plasmasphere. Scintillation activity in the nighttime equatorial anomaly regions of the low-latitude ionosphere, and in the high-latitude ionosphere, is known to peak at sunspot maximum, the next maximum being around 2000. Figure 3 is an example of GPS amplitude and phase scintillation indices S_4 and σ_ϕ recorded at the southern high-latitude ground station Casey on 23 September 1998. Hence the incidence of scintillation activity in the L band signals from GPS measured in both the occultation experiments and in the vertical direction is another area for study.

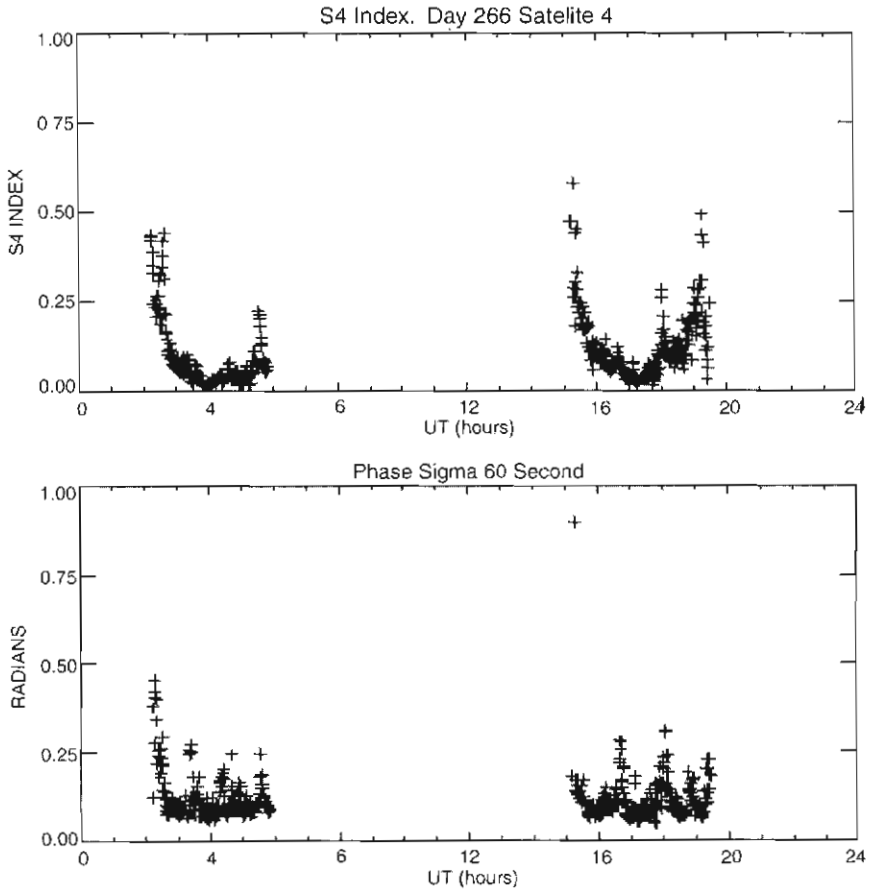


Figure 3. *GPS satellite number 4 amplitude and phase scintillation indices S_4 and σ_ϕ recorded at the southern high-latitude ground station Casey on 23 September 1998. The scintillation activity is calculated every minute.*

There is some indication that the occurrence of well-developed equatorial anomaly crests and their location is a precursor of the development of nighttime equatorial irregularities (Jayachandran *et al.*, 1997). The combination of both ground-based and satellite-based measurements would provide an excellent opportunity to further our understanding of the low-latitude ionosphere.

23.6 Conclusions

LEO microsatellites such as the planned Australian microsatellite FedSat provide the opportunity for researchers in Space Sciences to further investigate those regions of the ionosphere for which existing equipment and techniques are either difficult or impossible to apply. In the future, constellations of microsatellites may provide a global coverage of the ionosphere, previously not possible with ground-based equipment.

23.7 Acknowlegements

This work is supported by the Australian Cooperative Research Centre for Satellite Systems, and by an Australian Research Committee grant. C. McKinnon and N. Shilo are holders of Australian Postgraduate Awards and P. Webb and I. Horvath are holders of La Trobe University Postgraduate Awards.

References

Hajj, G.A., Ilbanez-Meir, R., Kursiniski, E.R. and Romans, L.J. (1994). Imaging the ionosphere with the Global Positioning System. *International Journal of Imaging Systems and Technology,* 5(2): 174–184.

Hajj, G.A. and Romans, L.J. (1998). Ionospheric electron density profiles obtained with the Global Positioning System: results from the GPS/MET experiment. *Radio Science,* 3: 157–190.

Horvath, I. and Essex, E.A. (2000). Using observations from the GPS and TOPEX satellites to investigate night-time TEC enhancements at mid-latitudes in the southern hemisphere during a low sunspot number period. *Journal of Atmospheric and Solar Terrestrial Physics,* 62 (5): 371–391.

Horvath, I. and Essex, E.A. (2001). GPS and TOPEX/Poseidon mid-latitude trough observations in the southern hemisphere at low sunspot numbers. In: Morris, R.J. and Wilkinson, P.J. (Eds.). *ANARE Reports 146.* Australian Antarctic Division. Pp. 307–319.

Jacobson, A.R., Hoogeveen, G., Carlos, R.C., Wu, G., Fejer, B.G. and Kelley, M.C. (1996). Observations of inner plasmasphere irregularities with a satellite-beacon radio-interferometer array. *Journal of Geophysical Research,* 101: 19665–19682.

Jayachandran, P.T., Sri Ram, P., Somayajulu, V.V. and Rama Rao, P.V.S. (1997). Effect of equatorial ionisation anomaly on the occurrence of spread-F. *Annales Geophysicae,* 15: 255–262.

Klobuchar, J.A., Doherty, P.H., Bailey, G.J. and Davies, K. (1994). Limitations in determining absolute total electron content from dual-frequency GPS group delay measurements. In: Kersley, L. (Ed.). *Proceedings of the International Beacon Satellite Symposium*. Aberystwyth, UK, 11–15 July. Pp. 1–4.

Raymund, T.D. (1994). Ionospheric tomography algorithms. *International Journal of Imaging Systems and Technology*, 5(2): 75–85.

Ruis, A., Ruffini, G. and Cucurull, L. (1997). Improving the vertical resolution of ionospheric tomography with GPS occultations. *Geophysical Research Letters*, 24: 2291–2294.

Titheridge, J.E. (1976). Plasmapause effects in the topside ionosphere. *Journal of Geophysical Research*, 81: 3227–3233.

Tyler, G.L. (1987). Radio propagation experiments in the outer solar system with Voyager. *Proceedings of IEEE* 75: 1404–1431.

Webb, P.A. and Essex, E.A. (1997). A simple model of the ionosphere plasmasphere system. In: Kulessa, A., James, G., Bateman, D. and Tobar, M. (Eds.). *Proceedings of the Workshop on Applications of Radio Science*, WARS'97. Barossa Valley, Australia, 21–23 September. Pp. 190–195.

Villani, A. and Essex, E.A. (1996). Computerised ionospheric tomography down the eastern coast of Australia. In: Goodman, J.M. (Ed.). *Proceedings of the Ionospheric Effects Symposium*. Alexandria, Virginia, USA, May. Pp. 19–24.

24. MODELLING THE PLASMASPHERE

P.A. Webb and E.A. Essex

Cooperative Research Centre for Satellite Systems
Department of Physics
La Trobe University
Bundoora Victoria 3083 Australia
(email: p.webb@latrobe.edu.au)

Abstract

With the recent advent of Global Positioning System (GPS) satellites as a research tool in studying the ionised regions that surround the Earth, a need has arisen for a simple yet accurate model for the plasmasphere, the region above the ionosphere. A model based on diffuse equilibrium has been developed, which is combined with the International Reference Ionosphere model to produce a global ionosphere-plasmasphere model. Some preliminary results from the model are presented and discussed.

24.1 Introduction

24.1.1 Effects of ionised mediums on signal propagation

The ionosphere is a layer of ionisation that surrounds the Earth, starting from an altitude of ~ 60 km and with a median altitude of approximately 400 km. The extension of the upper ionosphere along the Earth's closed magnetic field lines is known as the plasmasphere. A radio signal transmitted from a satellite to a ground receiver travels through this ionised medium; the effects on the propagating signal depend on the frequency and on the plasma density along the ray path. The effect of the plasma is measured in Total Electron Content (TEC), which is defined as the total number of electrons in a column with a cross sectional area 1 m^2 along the ray path. TEC is measured in Total Electron Content Units (TECU), where 1 TECU = 10^{16} electrons m^{-2}. By studying the TEC derived from satellite signals it is possible to study the ionised medium through which they have propagated.

24.1.2 Global positioning system satellites

The GPS satellites orbit at an altitude of 20 200 km with an orbital plane inclination of 55 degrees. Radio signals received from GPS satellites must propagate through some 19 500 km of the plasmasphere and the underlying ionosphere. The plasmasphere's electron density is on average one to four orders of magnitudes less than the ionosphere's. However, as a result of the greater distance travelled through the plasmasphere, the total amount of the plasmasphere's plasma traversed by the signal is usually appreciable compared with that in the ionosphere. Consequently the plasmasphere's effect on a GPS signal needs to be considered, especially around solar cycle minimum and at night.

24.1.3 La Trobe's electron density model

An increasing amount of the ionospheric electron density research being conducted at La Trobe University is based on GPS satellite data. It has therefore become important to understand the effect the plasmasphere's electron content has on GPS measurements, so that they can be correctly interpreted. To study this effect, a global electron density model of the ionosphere and plasmasphere has been developed.

In the past many computer-based models of the plasmasphere have been created, the FLIP model (Torr *et al.*, 1990) and the Sheffield model (Bailey and Balan, 1996) being two current examples. The primary aim of these models is to provide an understanding of, and to model the complex physical processes of the plasma flows along the magnetic field lines. In general these models required a computer workstation, because of the large number of calculations required. The primary aim of our research, however, is to produce a simple yet accurate global electron density model which can be run on a desktop personal computer (PC). The major goal of the model is to allow a user to specify a ray path between any two locations and have the model quickly and accurately determine the electron density at each point along the ray path. With this information, parameters such as TEC can easily be calculated.

24.2 An overview of the plasmasphere

The ionosphere is created by photoionisation of the neutral upper atmosphere by extreme ultraviolet and ultraviolet solar radiation. The ionosphere typically has a peak around 300 km altitude, which can vary between 200 km and 600 km. The peak ionospheric electron density is of the order 10^{10} to 10^{12} electrons m^{-3}. The peak height and the peak electron density depend on solar activity, geographical location, time and season.

In the 1950's, through studies of the propagation of electromagnetic waves along the Earth's magnetic field lines, it was found that the region beyond the ionosphere, now known as the plasmasphere, had unexpectedly high electron densities. It was realised that the source of this plasma was the reaction:

$$O^+ + H \Leftrightarrow O + H^+ \qquad (1)$$

which is accidentally energetically resonant and proceeds rapidly in both directions.

Chemical equilibrium as controlled by Equation (1) is established at altitudes where the oxygen ions (O^+) are the dominant species. Consequently, the lighter hydrogen ions (H^+) experience an outward force due to the ambipolar electric field created between the oxygen dominated ions and the electrons. Being charged, the H^+ are constrained to move parallel to the Earth's magnetic field lines. At lower latitudes where the Earth's field lines are closed, the upward flowing H^+ are trapped in the region centred on the magnetic equator. In the polar regions where the

field lines are open, the H^+ do not become trapped and flow out into space, producing what is called the 'polar wind'. The trapped plasma forms the plasmasphere, which generally extends out to three to six Earth radii (some 20 000 km to 40 000 km), with plasma densities of the order of 10^9 electrons m^{-3}. The shape and density of the plasmasphere is constantly changing, primarily due to changes in the underlying ionosphere (the source and sink of the plasmasphere) and the effects of magnetic storms.

24.3 An ionosphere-plasmasphere global electron density model

24.3.1 Overview of the model and its required inputs from existing models

The global electron density model of the ionosphere-plasmasphere that has been developed is based on diffusive equilibrium, with chemical equilibrium used at low altitudes. The model requires three outside major sources of data:

(i) Electron and ion temperatures.

(ii) The neutral temperature, oxygen and hydrogen densities.

(iii) Ionospheric electron density.

The required neutral parameters are obtained from the MSIS-90 model (Hedin, 1991, 1986). The ionospheric electron densities are obtained from the IRI95 model (Bilitza, 1995). Both these models can be down loaded as FORTRAN programs from the NSSDS Web Site (NSSDC, 1999). The electron and ion temperatures are discussed below.

24.3.2 Electron and ion temperatures

The electron and ion temperatures are calculated using a modified version of the upper ionosphere and plasmasphere temperature model published by Titheridge (1998). While it is beyond the scope of this paper to discuss these modifications in detail, in summary they are:

(i) The separation of the geographic and geomagnetic co-ordinates.

(ii) The modifications to the sunrise and sunset shaping parameters.

(iii) Re-calculation of the base temperature and gradient functions, using a reduced and modified satellite data set.

(iv) The modification of the base temperatures to remove effects of averaging over a seven hour time period during both day and night.

The reader should refer to Titheridge (1998) for a description of these terms and their relevance to the temperature model.

Two examples of electron temperature profiles obtained from several models are shown in Figure 1. The plots include profiles from the FLIP model, Titheridge's original model, the modified version of Titheridge's model, and the IRI95 model. Note that the IRI95 temperatures are not smoothly varying with altitude, compared with the other models, which

is due to the IRI95 profiles being calculated from independent satellite data sets (Brace and Theis, 1981).

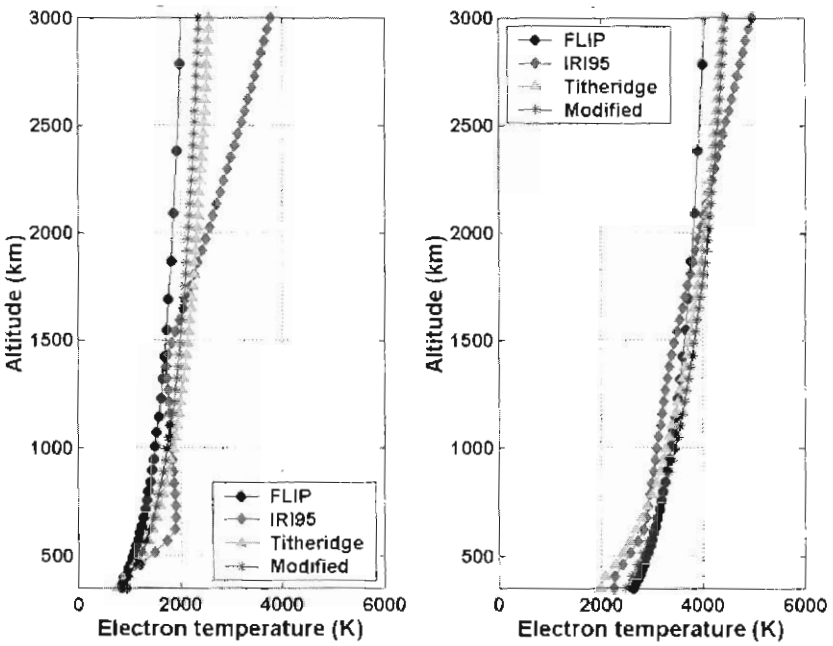


Figure 1. Comparison example between various models electron temperature output.

While there is generally good agreement between the different temperature models during the day, there are large differences at night. One of the reasons for this is the different bases of the respective models. The FLIP model is theoretical, the IRI95 is an empirical model, and the Titheridge model is a combination of theory and empirical data. This is a small example of one of the major problems with research in the upper ionosphere and plasmasphere, namely the lack of consistency between models and observed temperatures. A modified version of Titheridge's model is used because no other 'stand-alone' temperature model for this region exists (the IRI95 model does not extend above 3000 km altitude).

24.3.3 Diffusive equilibrium approach

The model that has been developed is based on diffusive equilibrium. Because the plasma is constrained to move along magnetic field lines, the diffusive profiles must be calculated along the field lines rather than simply in the vertical direction as is the case with the neutral atmosphere. The diffusive equilibrium equations used are given by

Titheridge (1972) and are reproduced in Equations (2) and (3). The scale height is given by:

$$H_j = \left(\frac{kT_t}{m_j g}\right)\left(1 - \frac{m_a T_e}{m_j T_t}\right) \tag{2}$$

where H_j is the j^{th} ion scale height, k is the Boltzmann constant, T_i is the ion temperature, m_j is the j^{th} ion mass, g is the gravitational acceleration, m_a is the mean ion mass, T_e is the electron temperature, and the total temperature is $T_t = T_i + T_e$. Taking n_{j0} as the ion number density at height h_0 where the total temperature T_{t0}, the ion number density n_j at the height $h_0 + \Delta h$ where the total temperature T_t is given by:

$$n_j T_t = n_{j0} T_{t0} \exp\left(-\frac{\Delta h}{H_j}\right) . \tag{3}$$

As Titheridge notes, 'the form given above is particularly suitable for numerical integration'. If the density of a given ion species at a certain base height is known, then by using incremental steps, its density can be calculated at some greater height. This is only possible if the temperatures are known at each step and no chemical ion production or loss is occurring. Because of the large amount of production and loss, diffusive equilibrium profiles are not always accurate in the ionosphere, and chemical equilibrium profiles need to be considered.

The simplest approach to producing an upper ionosphere-plasmasphere model is to consider O^+ and H^+ ions only. Since O^+ is the dominant species above ~ 150 km, its profiles can accurately be modelled using a diffusive profile at altitudes above the ionosphere's peak density. To model the H^+, chemical equilibrium is used in the ionosphere and diffusive equilibrium is used for the upper ionosphere and plasmasphere. The ion density profiles produced by this approach agree with the observed profiles when a magnetic field line (more correctly the magnetic flux tube) is 'saturated', which results in zero net H^+ flow between the ionosphere and the plasmasphere. The major reasons for this agreement is the approximate equality of the diffusion and chemical scale heights (Richards and Torr, 1985) for the H^+ when they are the minority ion species. The close equality of the scale heights alleviates the need to determine a boundary at which to convert from chemical to diffusive control of the H^+ profile, since in this case diffusive equilibrium can be used from the starting altitude.

However, the plasmasphere is rarely saturated, except at lower latitudes where the magnetic field line (flux tube) volumes are small, so changes in ion density initiate a quick response. This lack of saturation is primarily due to solar magnetic storms that periodically empty the outer plasmasphere, after which it starts to refill. Diurnal plasmaspheric variations also occur due to changes in ionospheric and plasmaspheric temperatures, changes in the ionospheric electron

densities and variations in the neutral atomic densities. The result is a highly dynamic plasmasphere with plasma consisting predominantly of H^+ flowing up into the plasmasphere or flowing down into the ionosphere. Unless the field line is saturated, a simple model that 'projects up' (meaning to use numerical integration in a series of small altitude 'steps') using chemical and diffusive equilibrium from the ionosphere will generally over estimate the true density in the plasmasphere by several orders of magnitude. This is because a field line that is in the process of refilling after a storm will have H^+ densities less than its saturated values, which are equal to the values obtained by projecting up the field line with chemical and diffusive equilibrium. This approach also has the problem that H^+ profiles projected up from northern and southern hemispheres along a given field line will rarely have the same densities at the magnetic equator, which in reality are the same.

24.3.4 Solutions to diffusive equilibrium problems

A solution to these problems was suggested by P.G. Richards (private communication, author of FLIP model mentioned previously) and forms the basic principle behind the developed model. Assuming that the total H^+ content of a field line is known, then to a good approximation the equatorial density of a field line is given by the total content divided by the total volume of the field line. This is due to most of the volume been centred on the magnetic equator, where the density is only slowly varying. This approximation is generally correct to within 5%. Starting from this equatorial density, the H^+ density profile is projected down the field line into both hemispheres using diffusive equilibrium (Equations (2) and (3)). This automatically solves the problem of 'equatorial mismatch' between the H^+ densities, which occurs when the profiles were projected up the field lines from the two hemispheres. A lower initial H^+ equatorial density, which occurs if the H^+ tube content is less than the saturated content, then represents a field line depleted by magnetic storm activity. The O^+ profile is still obtained by projecting up from the underlying ionosphere. This is appropriate since the ionospheric O^+ densities are primarily controlled by chemical processes and the solar flux interactions with the neutral atmosphere, and consequently recover very quickly after a magnetic storm.

This approach creates the problem that as the diffusive H^+ profiles reach down into the ionosphere they need to be matched with the chemical equilibrium H^+ profiles. A desire to have close agreement between the model under development and the FLIP model gives some guidance as to the desired profile that the H^+ should have in this region. A method has been developed which smoothly joins the two profiles and is in good agreement with the FLIP model. While the details are also beyond the scope this paper, the method is based upon an equation derived by Richards and Torr (1985), which estimates the height where chemical loss is equal to diffusive loss. The corrected equation (Rasmussen *et al.*, 1993) is given below:

$$z_0 = z_r - \left(\frac{H_1 H_2}{H_1 + H_2}\right) \ln\left(\frac{7.5x10^{19} T_i^2}{H^2 n_r(O^+) n_r(O)}\right) \quad (4)$$

where z_0 is the height where chemical and diffusive heights are equal, z_r is a reference height, H is the diffusion scale height, H_1 is the O^+ scale height, H_2 is the O scale height, $n_r(O^+)$ is the O^+ number density and $n_r(O)$ is the O number density, both at the reference height r. All heights are in km, and densities are in m^{-3}. Based on this equation a weighting function has been derived which allows the two hydrogen ion profiles to be smoothly joined.

24.3.5 Time evolution of the field line

A procedure for the accurate production of the H^+ profiles along a given field line has been presented above. The final consideration is the time evolution of the H^+ content of the field line. As indicated earlier, the source of H^+ in the plasmasphere is represented in Equation (1). The production rate of H^+ in m^{-3} s^{-1} is given by (see for example Richards and Torr, 1985):

$$P = 2.5x10^{-17} T_n^{1/2} n(H) n(O^+) \quad (5)$$

and the loss rate by:

$$L = 2.2x10^{-17} T_i^{1/2} n(O) n(H^+) \quad (6)$$

where T_n is the neutral temperature, and n(#) is the number density of the given ion or neutral atom species (denoted by #). Most of the H^+ production and loss occurs in the upper ionosphere and lower plasmasphere, due to the rapid reduction in the neutral oxygen and oxygen ion densities with altitude, which retards both production and loss at higher altitudes. Since at any time the neutral densities and ion densities are known, both the production and loss of H^+ for a given field line can be calculated. Assuming the density profiles do not change over a small time increment, the net change in the H^+ content of the field line can be calculated. Advancing time by the given time increment, a new equatorial H^+ density can be calculated as before and the process repeated. This allows the time development of the H^+ content of a magnetic field line to be modelled.

The other dynamic aspect of the plasmasphere that needs to be considered is the motion through space of the individual magnetic field lines, or more correctly the motion of the plasma aligned with a magnetic field line. At low magnetic latitudes the plasma remains attached to the field lines and co-rotates with the Earth. At higher latitudes the electric fields in the Earth's magnetosphere cause the plasma to move from one magnetic field line to another, with the magnitude of this motion being directly related to solar activity. During periods of large solar magnetic storms the outer plasmasphere is torn away from the Earth by the increased strengths of these electric fields and lost into the outer magnetosphere.

In the first version of the model under development it had been assumed that the plasma co-rotates with the Earth magnetic field lines at all latitudes. The effects of magnetic storms are included in this case by assuming that a field line is emptied of its plasma if it is between 1200 magnetic local time (MLT) and 1800 MLT, and is currently outside the plasmapause, the outer boundary of the plasmasphere. Ignoring the plasmasphere's experimentally observed evening bulge, the location of the plasmapause (the outer plasmasphere's boundary) in the magnetic equatorial plane is determined by using Equation (7) (Gallagher *et al.*, 1995), where L_{pp} is in Earth radii.

$$L_{pp} = 5.6 - 0.46K_p. \tag{7}$$

Equation (7) shows that as the magnetic storm activity increases (higher K_p) the plasmapause moves inwards. The choice of the time interval 1200 MLT to 1800 MLT for the flux tubes to be emptied is a first order approximation to the results obtained when a magnetospheric electric field is included, which is the next level of sophistication in modelling this process. At other times the plasma generally co-rotates with the Earth, even if outside the plasmapause.

24.3.6 Calculating electron densities

Up to this point, this paper has discussed a general method to obtain the H^+ and O^+ densities along a field line. However, many observations such as TEC require the electron density and not the ion densities. This is obtained by assuming the plasmasphere and ionosphere has a neutral charge at each point, thus the electron density is taken to be equal to the sum of the H^+ and O^+ densities at any given point. The electron density at a required point can be determined by calculating the electron densities along the field line that the point of interest lies, as described previously. The field line is broken up into pre-determined distance intervals and interpolation is used to determine the electron density at the required point of interest. The electron density along a ray path is obtained by breaking the path into a series of points, at which the electron densities are separately calculated using interpolation along their respective field lines.

24.4 Examples of results from the model

Because the model has only been recently completed, the results that will be presented here are preliminary. They do however indicate the type of results that the model is capable of producing, noting that its primary goal is to calculate the electron density at any point surrounding the Earth up to a radius of five Earth radii.

24.4.1 Total electron content

With the ability to calculate electron densities along a ray path, it is straightforward to calculate TEC. Figure 2 shows an example where the vertical TEC has been calculated between the north and south magnetic poles at 150° magnetic longitude.

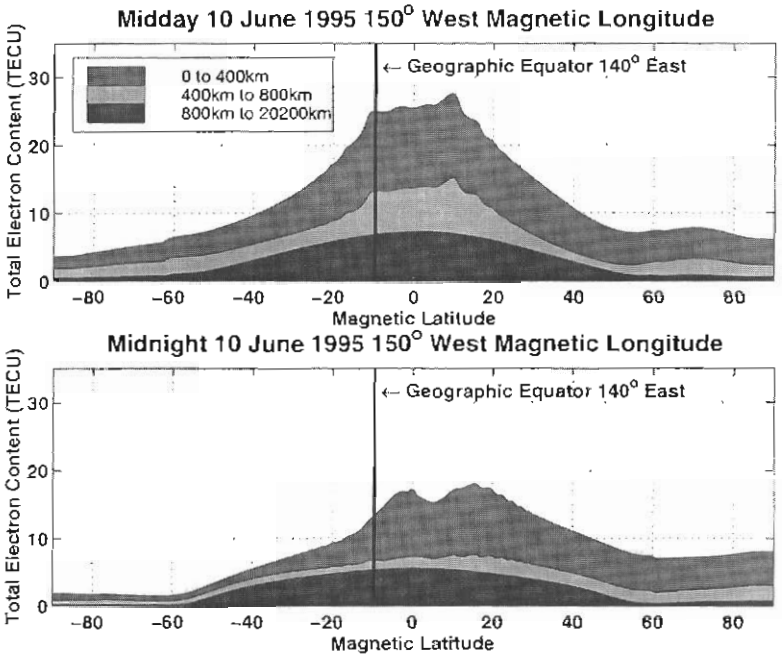


Figure 2. *Example of vertical TEC calculation.*

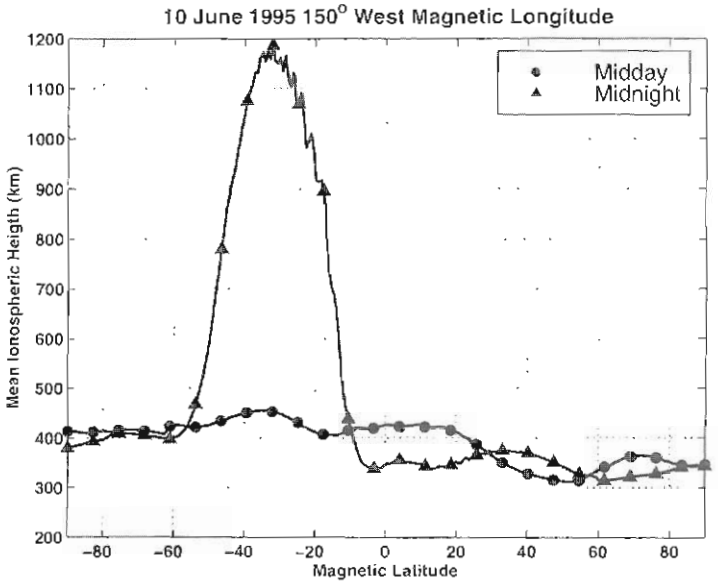


Figure 3. *Example of median ionospheric height.*

Unlike direct satellite observations where only the total TEC can be measured, the contributions to the TEC from various altitude bands can be calculated. The examples shown in Figure 2 are for near solar minimum, at midday and midnight respectively. The height of 20 200 km is chosen as it corresponds to the altitude of the GPS satellites. It should be noted that the TEC above 20 200 km predicted by the model is at the most 0.15 TECU, which is insignificant.

Taking 800 km as being the transition from the upper ionosphere to the plasmasphere, the following features can be noted. Figure 2 clearly shows the symmetrical nature of the plasmasphere resulting from the magnetic field lines that control its shape. This leads to the offset from the geographic equator of the plasmasphere and the ionospheric anomaly by 10° in this case. The plasmasphere exists mainly between magnetic latitudes of ±50°, and hence there is normally little contribution to the TEC from above 800 km in the polar regions. In the equatorial region, the plasmasphere contribution to the TEC is the order of 30%. Note that during the night in the southern hemisphere this contribution is over 50% and one consequence of this will be discussed below.

24.4.2 Median ionospheric height

Knowledge of the electron densities along a ray path allows another ionospheric parameter, the Median Ionospheric Height to be calculated. This is the altitude at which half the TEC is above and half is below along a ray path. It is required for example when converting from slant TEC to vertical TEC, and is normally assumed to be a constant altitude of 400 km. In Figure 3 the Median Ionospheric Height has been calculated along the same magnetic longitude as Figure 2.

Figure 3 shows that a constant Median Ionospheric Height of 400 km, is in this case generally a good approximation, except during the night at southern middle-latitudes where the correct value is over 1000 km. This is caused by the plasmasphere contributing approximately 50% to the TEC, as shown in Figure 2. Consequently any vertical TEC calculated from slant TEC using the 400 km approximation will be in error due to the actual incident angle of the ray path of the satellite signal being slightly smaller than the assumed incident angle.

24.4.3 Further results

Other results that can be obtained from the model include studying the plasmasphere's recovery after magnetic storms, and the effects on observed TEC measurements from GPS satellites. The length of time taken by the plasmasphere to recover can be estimated by comparing the refilling of the magnetic field lines at various times during the recovery phase. With the ability to calculate the plasmaspheric contribution to the observed TEC from a GPS satellite, it is possible to remove it so that the ionospheric contribution can be separately determined. This correction will allow more accurate studies of the ionosphere to be undertaken using GPS satellite signals. The model will be used to study, for example,

the expected TEC along ray paths between two satellites, such as the Australian FedSat satellite and the GPS constellation.

24.5 Conclusion

A global electron density model of the ionosphere-plasmasphere has been developed at La Trobe University, which allows the calculation of electron densities at any point up to five Earth radii. This allows the densities along a given ray path to be calculated, from which parameters such as the Total Electron Content and Median Ionospheric Height can be determined. Knowledge of these parameters is important in allowing the correct interpretation of GPS satellite signals received by ground stations after passing through the ionosphere and plasmasphere.

24.6 Acknowledgements

This research is supported by an Australian Research Council (ARC) grant. P. Webb is supported by a La Trobe University Postgraduate Award (LUPA) and the Cooperative Research Centre for Satellite Systems (CRCSS). The authors wish to thank P.G. Richards of the University of Alabama and J. Titheridge from The University of Auckland for advice in the development of the model.

References

Bailey, G.J. and Balan, N. (1996). A low-latitude ionosphere-plasmasphere model. In: Schunk, R.W. (Ed.). Solar-Terrestrial Energy Program: Handbook of Ionospheric Models. Pp. 173–206.

Beggs, H.M., Essex, E.A. and Rasch, D. (1994). Antarctic polar cap total electron content observations from Casey Station. *Journal of Atmospheric and Terrestrial Physics*, 26: 659–666.

Bilitza, D. (1995). International Reference Ionosphere. http://nssdc.gsfc.nasa.gov/space/model/ionos/iri.html.

Brace, L.H. and Theis, R.F. (1981). Global empirical models of the ionospheric electron temperature in the upper F-region and plasmasphere based on in situ measurements from the Atmospheric Explorer-C, ISIS-1 and ISIS-2 satellites. *Journal of Atmospheric and Terrestrial Physics*, 43 (12): 1317–1343.

Breed, A.M. (1992). Ionospheric Total Electron Content studies using satellite radio transmissions. Master of Applied Science Thesis. University of South Australia.

Brinton, H.C., Grebowsky, J.M. and Brace, L.H. (1978). The high latitude F-region at 300 km: thermal plasma observations from AE-C. *Journal of Geophysical Research*, 83: 4767–4776.

Crowley, G., Carlson, H.C., Basu, S., Denig, W.F., Buchau, J. and Reinisch, B.W. (1993). The Dynamic ionospheric polar hole. *Radio Science*, 28: 401–413.

Daniell, R.E. Jr., Brown, L.D., Anderson, D.N., Fox, M.W., Doherty, P.H., Decker, D.T., Sojka, J.J. and Schunk, R.W. (1995). Parameterized ionospheric model: a global ionospheric parameterization based on first principles models. *Radio Science*, 30: 1499–1510.

Gallagher, D.L., Craven, P.D., Comfort, R.H. and Moore, T.E. (1995). On the azimuthal variation of core plasma in the equatorial magnetosphere. *Journal of Geophysical Research*, 100 (A12): 23597–23605.

Hedin, A.E. (1986). MSIS-86 thermospheric model. *Journal of Geophysical Research*, 92 (A5): 4649–4662.

Hedin, A.E. (1991). Extension of the MSIS Thermosphere Model into the middle and lower atmosphere. *Journal of Geophysical Research*, 96 (A2): 1159–1172.

Kelley, M.C. (1989). The Earth's Ionosphere: Plasma Physics and Electrodynamics. San Diego Academic Press. Pp. 326–329.

Kersley, L., Pryse, S.E. and Wheadon, N.S. (1988). Amplitude and phase scintillation at high latitudes over northern Europe. *Radio Science*, 23: 320–330.

NSSDC (1999). Space Physics at the National Space Science Data Centre http://nssdc.gsfc.nasa.gov/space/space_physics_home.html.

Rasmussen, C.E., Guiter, S.M. and Thomas, S.G. (1993). A two-dimensional model of the plasmasphere: refilling time constants. *Planetary and Space Science*, 41 (1): 35–43.

Richards, P.G. and Torr, D.G. (1985). Seasonal, diurnal, and solar cyclical variations of the limiting H^+ in the Earth's topside ionosphere. *Journal of Geophysical Research*, 90 (A6): 5261–5268.

Rodger, A.S. and Graham, A.C. (1996). Diurnal and seasonal occurrence of polar patches. *Annales Geophysicales*, 14: 533–537.

Titheridge, J.E. (1972). Determination of ionospheric electron content from the Faraday rotation of geostationary satellite signals. *Planetary and Space Science*, 20: 353–369.

Titheridge, J.E. (1998). Temperatures in the upper ionosphere and plasmasphere. *Journal of Geophysical Research*, 103 A2: 2261–2277.

Torr, M.R., Torr, D.G., Richards, P.G. and Yung, S.P. (1990). Mid- and low-latitude model of thermospheric emissions. *Journal of Geophysical Research*, 95 (A12): 21147–21168.

Weber, E.J., Buchau, J., Moore, J.G., Sharber, J.R., Livingstone, R.C., Winningham, J.D. and Reinisch, B.W. (1984). F-layer ionization patches in the polar cap. *Journal of Geophysical Research*, 89: 1683–1694.

25. CONJUGACY OF PC3-4 WAVES IN THE HIGH-LATITUDE MAGNETOSPHERE

T.A. Howard[1], F.W. Menk[1] and R.J. Morris[2]

[1] Department of Physics
Cooperative Research Centre for Satellite Systems
University of Newcastle
Callaghan NSW 2308 Australia

[2] Atmospheric and Space Physics
Australian Antarctic Division
Kingston Tasmania 7050 Australia

Abstract

Pc3-4 pulsations observed at high-latitudes have been examined for evidence of magnetic field line resonances. Previous studies have suggested that these pulsations are due to harmonics of the lower frequency Pc5 pulsations. We investigated near-conjugate properties of high-latitude Pc3-4 pulsations using magnetometers of the IMAGE array in northern Scandinavia and the fluxgate magnetometer at Davis station, Antarctica. Analysis included comparison of time series, dynamic power spectra, and cross-phase, cross-power and coherence properties between the conjugate points. Our analysis of 11 events from March, 1996 suggests that the Pc3-4 pulsations are not standing field line oscillations or harmonics of field line resonances. We discuss two possible mechanisms for their generation; the forced oscillation of field lines and propagation direct to the ionosphere. In both cases the pulsation source is the upstream ion-cyclotron resonance mechanism.

25.1 Introduction

This paper will focus on low frequency Pc3-4 (f ~ 10–50 mHz) waves in the high-latitude regions. At low-latitudes (L ~ 2–3), Pc3-4 pulsations are attributed to the standing oscillation of magnetic field lines at their eigenfrequency, called field line resonance or FLR (Ansari and Fraser, 1986; Menk *et al.*, 1994). However, at higher latitudes (L ≥ 6) the field lines are longer, and the resonant frequency moves into the lower frequency (1–10 mHz) Pc5 range (Samson and Rostoker, 1972; Waters *et al.*, 1995). Throughout local daytime, there is a largely continuous Pc5 band observed in the high-latitude regions which is attributed to field line resonances (Waters *et al.*, 1995; Ziesolleck and McDiarmid, 1995). Pc3-4 signals are also often observed in the dayside magnetosphere.

Previous workers have identified high-latitude Pc3-4 as higher harmonics of Pc5 (Fukunishi and Lanzerotti, 1974a; Tonegawa and Fukunishi, 1984;

Ziessolleck *et al.*, 1997). They proposed that energy from an external source couples with Pc5 waves, which are nearly always present in the dayside magnetosphere to drive higher harmonics of Pc5 FLRs with frequencies in the Pc3-4 range. Signals of this nature would be expected to possess characteristics of a FLR. Such characteristics include a phase reversal in phase-latitude profiles and an amplitude peak in amplitude-latitude profiles (Waters *et al.*, 1991b; Howard and Menk, 2001), in addition to the conjugate point properties discussed later. Another possible Pc3-4 generation mechanism is that of modulated electron precipitation (Olson and Szuberla, 1997).

The use of conjugate point measurements to determine various properties of waves has been known for many decades, since Alfvén (1950) developed the stretched string analogy of magnetic field lines. A magnetic field line has its 'fixed' ends at conjugate ionospheres in opposite hemispheres. At low-latitudes the field lines are approximately dipolar and conjugate points are thus calculated using a dipole model. However, at high-latitudes the field becomes distorted and models such as the corrected geomagnetic (CGM, Gustafsson, 1992) or the Tsyganenko (1989) model must be used to determine conjugate point locations. The conjugate stations used in this study are the Longyearbyen-Davis pair. These, as shown in Figure 2, are not exact conjugates of each other. Other examples include the Mawson-Jan Mayen and Syowa-Husafell (Kato *et al.*, 1994) pairs.

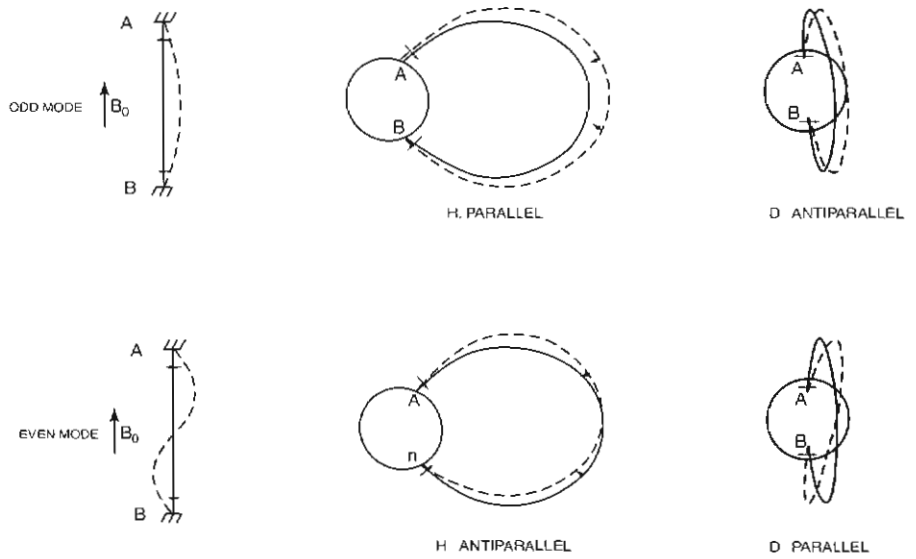


Figure 1. *The symmetry relations at magnetically conjugate points for the oscillation of the lines of magnetic force (Sugiura and Wilson, 1964). H and D represent the horizontal and easterly components respectively and the arrow indicates the direction of the magnetic perturbation.*

The first mathematical analysis of conjugate points was made by Sugiura and Wilson (1964). They showed that if standing field line oscillations are present, then at conjugate points one would detect H components which are in phase and D components which are out of phase for an odd mode (Figure 1) and vice versa for an even mode. Other wave features such as polarisation can also be determined by studying magnetic footprints at conjugate sites. Experimental studies by Lanzerotti *et al.* (1972) and Fukunishi and Lanzerotti (1974b) found that Pc3-4 activity around L = 4 is predominantly odd mode suggesting that the pulsations are caused by field line oscillations, possibly resonances.

This paper will focus on the conjugate point properties of Pc3-4 waves at L ≥ 10 in order to investigate whether the pulsations are higher

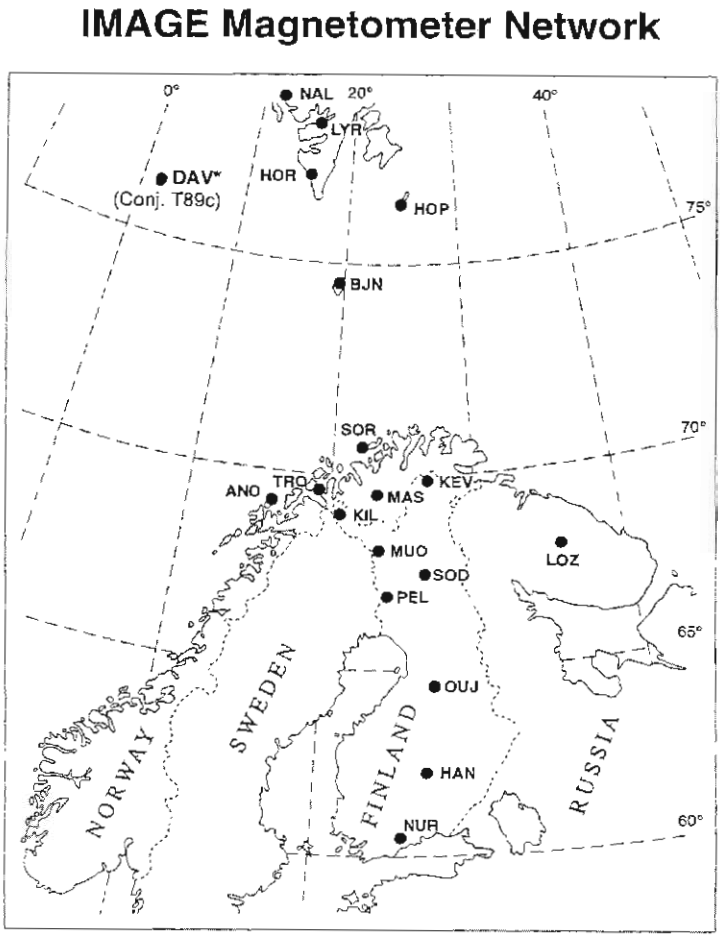


Figure 2. *The IMAGE array with the Davis conjugate point. The conjugate point was determined using the Tsyganenko (T89c) model.*

harmonics of Pc5. A previous study has examined the phase and coherence of high-latitude Pc3-4 and suggested the pulsations are not FLR harmonics (Howard and Menk, 2001). Further investigation is clearly required.

25.2 Data collection and analysis

Data were obtained using the IMAGE magnetometer array (CGM latitudes = 60°–79°) in northern Scandinavia and the magnetometer at Davis station (CGM latitude = -69°). IMAGE consists of 19 triaxial fluxgate magnetometers which sample in units of nT at 0.1 Hz with a resolution of 0.1 nT. Details on IMAGE can be obtained from the following web site: http://www.geo.fmi.fi/image/index.html. southern hemisphere data were from the fluxgate at Davis (details at http://www.aad.gov.au). For the time interval chosen, the Davis magnetometer also sampled at 0.1 Hz. Figure 2 shows the IMAGE array along with the location of the Davis conjugate point calculated with the Tsyganenko T89c model. This is a semi-empirical model which takes the time of day, season and K_p value into account. The CGM model is basically a distorted dipole field model. Any CGM value in this paper is for an altitude of 100 km above sea level and epoch 1996. The Davis conjugate footprints for both models essentially map to the same location and are about 400 km south-west of Longyearbyen.

Data were obtained from the IMAGE and Australian Antarctic Division webpages. At all stations, data are recorded in geographic (X, Y, Z) coordinates and a coordinate rotation into geomagnetic (H, D, Z) coordinates was performed. For this, an accurate determination (in geographic latitude and longitude) of the geomagnetic pole was needed. This can be determined experimentally, such as the alignment with the compass needle; theoretically, such as from models such as CGM; or semi-empirically, such as with the Tsyganenko model. For the purposes of this study, the experimentally determined value was used (geographic latitude -64.7°, longitude 138.6°; http://www.ngdc.noaa.gov/seg/potfld/faqgeom.html). From here, spherical geometry was used to determine the bearing of this pole from geographic north, which is compared with the value of D (declination angle of the magnetic field), obtained from: http://nssdc.gsfc.nasa.gov/space/cgm/cgm.html.

The coordinates were then rotated by the value of D (~ 78° for Davis). As the Z component always points 'downward' this component and the X component were rotated by a further 180° to accommodate for opposite hemispheres. All data were recorded in UT, where LT ~ UT +1 for IMAGE and LT ~ UT +2 for Davis.

Eleven events were selected from two weeks of data from March 1996, embracing one week of quiet and the other of disturbed magnetic activity. An event was selected if Pc3-4 was seen across most of the IMAGE array and was clearly resolved from noise using whole-day dynamic cross-phase and power spectra. For each event the time series were compared for each IMAGE station and at Davis. Next, dynamic

power spectra were examined for the Longyearbyen-Davis pair. Finally, the cross-phase, cross-power and coherence for this pair were determined. Both H and D components were considered for each measurement. The dynamic power, cross-power and cross-phase spectra were computed with a 256-point (43 minutes long) Fast Fourier Transform (FFT) stepping every 35 points and were used for event selection. Interstation phase was measured from individual 20 minute long cross-phase spectra (Howard and Menk, 2001). The dynamic spectra were weighted by $f^{1.5}$ to enhance the higher frequency features; spectral resolution was 0.83 mHz. For the dynamic power spectra, windowing was performed in the frequency domain using a truncated exponential taper stepping every five points, and a Hanning window of the same dimension as the FFT length was used for the cross-phase measurements.

Table 1. *List of events selected for analysis.*

Date (1996)	Frequency (mHz)	K_p	Start Time (UT)	Coherence Length (Howard & Menk 2001)
March 2	17	0	1030	$(1.5 \pm 0.6) \times 10^3$ km
March 2	17	0	1100	$(1.3 \pm 0.9) \times 10^3$ km
March 4	42	3-	0715	$(1.1 \pm 0.4) \times 10^3$ km
March 5	27	2+	0430	$(0.6 \pm 1.1) \times 10^3$ km
March 5	26	2-	0630	$(0.6 \pm 1.2) \times 10^3$ km
March 5	23	2-	1145	$(0.6 \pm 1.1) \times 10^3$ km
March 6	34	1	0800	$(0.8 \pm 0.9) \times 10^3$ km
March 19	24	4-	0845	$(1.1 \pm 0.6) \times 10^3$ km
March 23	27	2-	0730	$(1.4 \pm 0.7) \times 10^3$ km
March 23	21	1	1130	$(0.5 \pm 2.6) \times 10^3$ km
March 26	21	1+	1115	$(1.8 \pm 0.6) \times 10^3$ km

25.3 Results

A list of the events is shown in Table 1. Also included is the coherence length as determined by Howard and Menk (2001). Coherence length is the ground distance covered by a signal before its coherence, relative to a central reference station, falls below 0.65. Figure 3 shows time series data for several IMAGE stations and Davis for a typical event on 23 March 1996. The event had a frequency of 27 mHz and occurred around 0730 UT. Filter bandwidth was 20–35 mHz. For this example the H component results are shown. In all cases the H component was considerably larger than the D component but a signal could still be seen at the IMAGE stations for both components. Notice the packet structure of the wave in the top four plots from around 0735 to 0748 UT. The

packet structure for the signal at Davis does not resemble the structure at the other stations.

Figure 4 shows whole-day dynamic power spectra from Longyearbyen (LYR) and Davis (DAV) for the event shown in Figure 3. Note the almost continuous band of low frequency Pc5 activity at both stations up until ~ 1100 UT. The event appears as a dark patch around 0730 UT and 27 mHz in the LYR spectrum. However, this event has lower power in the DAV spectrum, where instead a later event is more prominent. Note the broadband bursts at both stations commencing at 0105, 1800 and 2155 UT. These are substorm events which are simultaneous everywhere. The other vertical features in the DAV spectrum beginning at 1115 and 1545 UT are due to local noise.

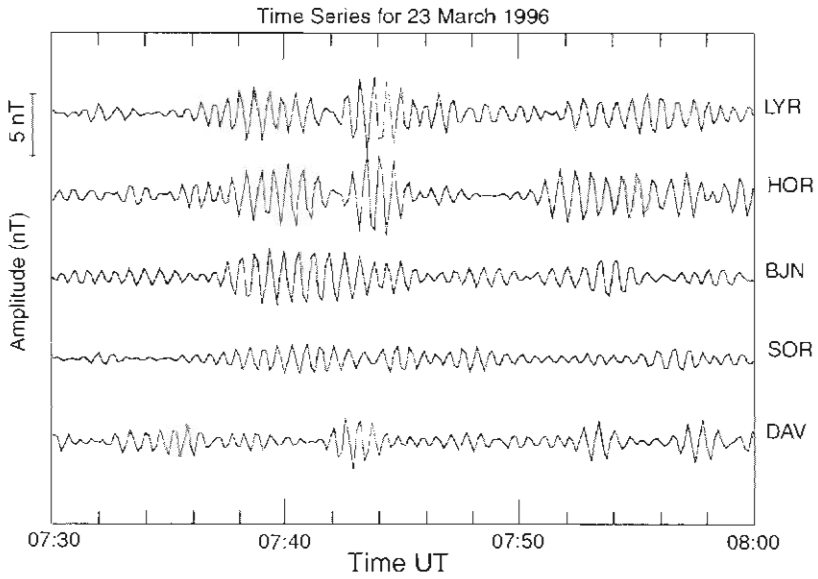


Figure 3. *Time series for the event on 23 March 1996. The time series is shown for Longyearbyen (LYR: Latitude = 78.2° CGM), Hornsund (HOR: 77.0° CGM), Bear Island (BJN: 74.5° CGM), Sørøya (SOR: 70.5° CGM) and Davis (DAV: -68.6° CGM). The signal is bandpass filtered at 20–35 mHz.*

Figure 5 illustrates how determinations of cross-phase and coherence are made, and includes the cross-power to assist in signal identification. The H components are shown, high pass filtered at 5 mHz. The top panel is the unfiltered time series for each station. Note the peak in cross-power around 27 mHz in Panel 2. This identifies the same frequency signal in both hemispheres. Note also the low coherence at this frequency (Panel 3), ~ 0.1, far too low to be considered reliable. This suggests the same Pc3 wavepacket is not seen at DAV and LYR.

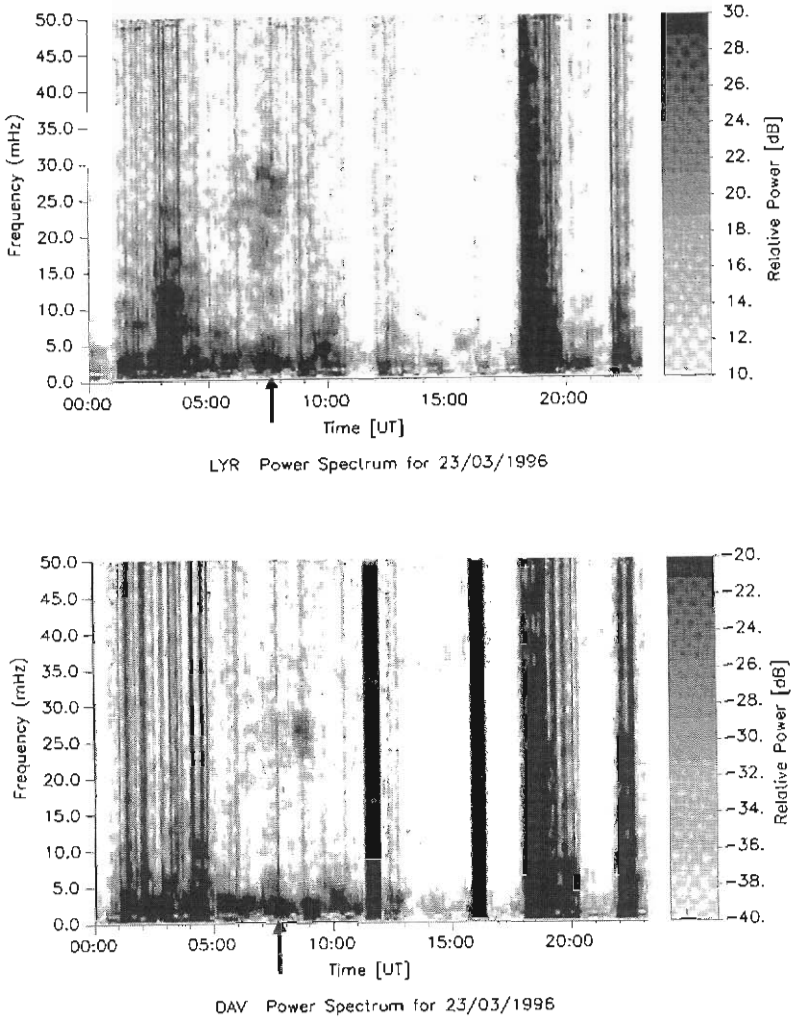


Figure 4. *Dynamic power series for the Longyearbyen-Davis conjugate pair, again for the event on 23 March 1996. The grey scale represents power intensity in dB as indicated in the bar to the right. The event, indicated by the arrow occurs at around 0730 UT with a frequency of around 27 mHz. The broadband signals detected at both stations after 1800 UT are due to substorm activity in the night sector.*

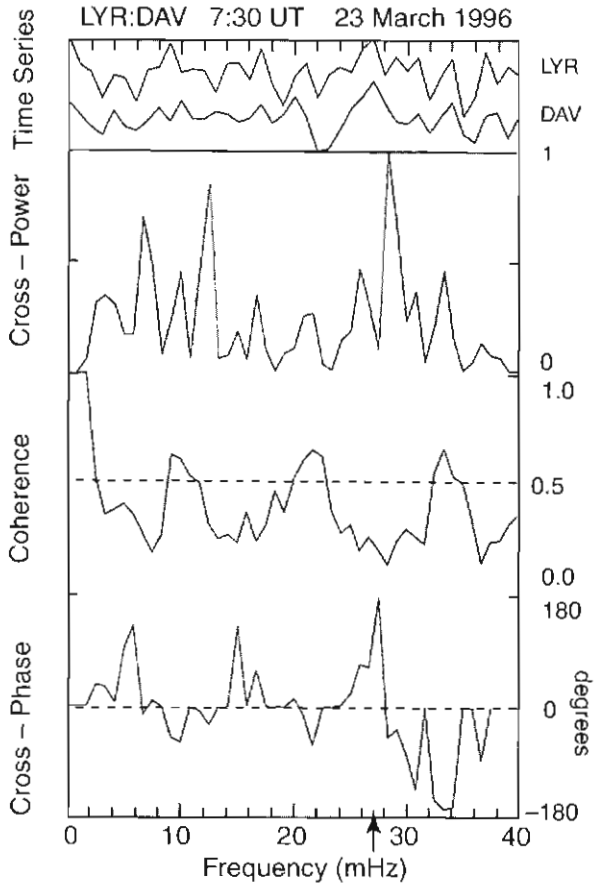


Figure 5. *Panel 1: Unfiltered time series for the LYR-DAV pair;*
Panel 2: Normalised relative cross-power spectra;
Panel 3: Relative coherence values;
Panel 4: Cross-phase values for the event on 23 March 1996. The H component is shown.

Figure 6 shows the same event, presented in Figures 4 and 5, over a different time scale and with a filter bandwidth of 1–4 mHz to show the Pc5 band present. It is clear that similar Pc5 pulsations were present at both locations, and analysis of this signal reveal a peak in coherence and cross-power around 2.5 mHz.

Of the eleven events examined here, none of the IMAGE data revealed any high values for coherence, or displayed similar time series, when compared with Davis data. For three events, interhemispheric peaks in cross-power were detected for one component, and for four others, values of coherence reached 0.5, but never was this the case for both H and D components.

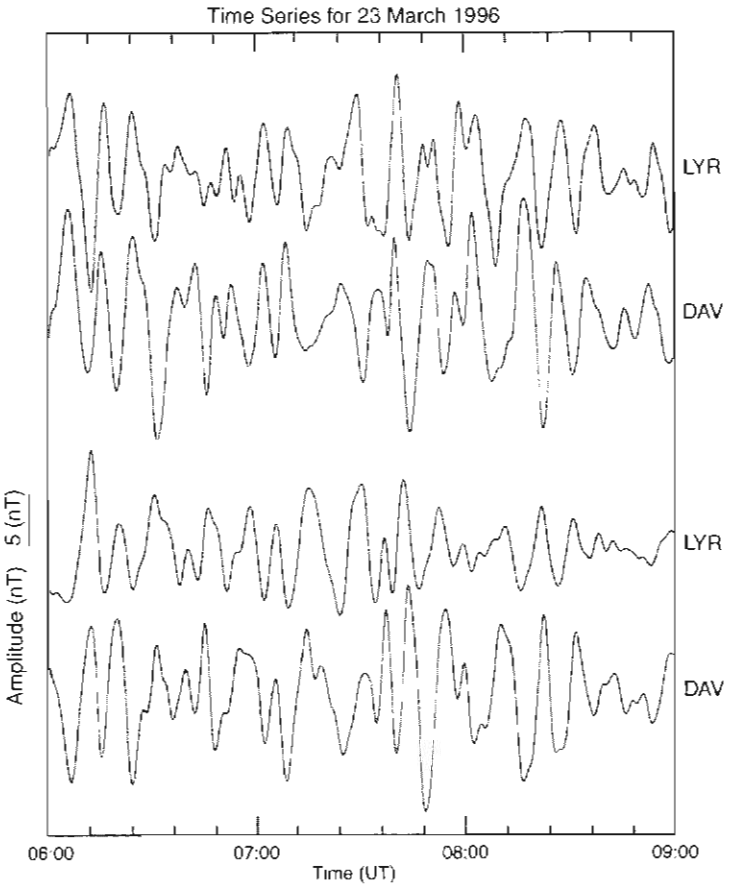


Figure 6. *Low frequency time series for H (top traces) and D (bottom traces) components for the same 23 March 1996 event. This has a filter bandwidth of 1–4 mHz revealing the Pc5 band at each station. Scaling of this Davis record is approximate.*

25.4 Discussion

For each Pc3-4 event the continuous Pc5 band was observed at the near conjugate points as shown in Figures 4 and 6. This indicates the presence of a Pc5 field line resonance (fundamental standing odd mode field line oscillation) at these latitudes and can also be used as an indication of closed field lines (Ables *et al.*, 1998). Pc3-4 pulsations were clearly visible across each station in the IMAGE array and at the conjugate point, but the packet structure was dissimilar and coherence was very low, at the conjugate station. Even in the cases where the coherence reached the threshold of acceptability (0.65), in no cases was this achieved in both H and D components.

The same Pc3-4 events exhibit high coherence over extended distances ($\geq 10^3$ km) as determined using the IMAGE stations (Howard and Menk, 2001). The distance between the actual conjugate point of Davis and Longyearbyen is ~ 400 km and falls well within the coherence length of each signal at IMAGE.

Summarising, the results of this study have shown that:

(i) Pc3-4 pulsations are observed simultaneously at near-conjugate high-latitude locations although the waveforms and spectral structure are different and the interhemispheric coherence is low.

(ii) The Pc3-4 pulsation events exhibit similar waveforms and spectral structure in the same hemisphere, measured using the IMAGE array. They maintain high coherence over ground distances typically exceeding 1000 km.

(iii) These results and the amplitude-latitude and phase-latitude results reported by Howard and Menk (2001), indicate that the Pc3-4 pulsations are not standing oscillations of field lines and do not exhibit properties of harmonics of field line resonances.

(iv) At the same time, sustained low frequency Pc5 pulsations are seen in both hemispheres. Their conjugate characteristics are consistent with standing fundamental field line resonance oscillations.

It is generally believed that the energy source for Pc3-4 signals is the ion-cyclotron instability involving backstreaming protons upstream of the solar wind (Greenstadt *et al.*, 1983). The resultant fast mode waves may penetrate the magnetosheath and enter the outer magnetosphere. Olson and Szuberla (1997) found short coherence lengths ($\leq$ 200 km) for burst-like Pc3-4 signals at high-latitudes and these are most likely connected with electron precipitation which is modulated by the Pc3-4 waves near the last closed field lines. This perturbs local ionospheric current systems, resulting in a broadband spectrum (extending into the Pc3 range) at the ground. This mechanism cannot account for the long-range coherent signals investigated in this study.

We consider two possibilities to explain the observations. We have previously shown (Howard and Menk, 2001) that the Pc3-4 events examined are most likely closely connected with properties of the solar wind. Accordingly, we believe that the pulsations' source mechanism is the upstream ion-cyclotron instability. Incoming fast mode waves may couple to field lines forcing them to oscillate with the same frequency as the incoming wave. As the wave propagates inward it forces further field lines to oscillate until all of its energy is transferred into the field. As the travel time for waves along field lines to the ground is longer for the outer field lines the waves would arrive at the lower latitudes first, thus creating the appearance of poleward propagation as found by Howard and Menk (2001). The amplitude of the signal detected on the ground would also appear to decrease with decreasing latitude which also agrees with the findings of Howard and Menk (2001), and with Bol'shakova and Troitskaya (1984). The lack of conjugacy found in this paper puts some constraints on this model.

Alternatively, the fast mode waves may propagate directly to the ionosphere, undergoing refraction though the magnetosphere as they did so. This model also predicts the direction of propagation determined by Howard and Menk (2001) and can account for the lack of conjugacy found here.

Figure 6 displays Pc5 properties between Longyearbyen and Davis. In this figure we have assumed a typical azimuthal wavenumber of m = -3 (Ziessolleck *et al.*, 1997; Mathie *et al.*, 1999) and adjusted the signal phase at Davis accordingly to account for the ~ 400 km separation between the Davis conjugate and Longyearbyen. The conjugate signals are seen to be in phase, while the D component is variable but generally around 90°–180° out of phase, as expected for odd mode Pc5 resonances. The exact phase relationship depends on the difference in latitude and longitude between the conjugate locations. Future work will also examine data from the January Mayen-Mawson pair, which are more nearly conjugate than Longyearbyen-Davis. Nevertheless, it is clear that the Pc5 pulsations are being observed on closed field lines, and that the Pc3-4 pulsations are not coherent at the conjugate locations.

25.5 Acknowlegements

We thank all members of the IMAGE team and those at ANARE responsible for operating the magnetometers and providing the data. TAH was supported by an Australian Research Council fellowship, and this work was supported by the Australian Research Council, Antarctic Science Advisory Committee, the University of Newcastle and the Cooperative Research Centre for Satellite Systems.

References

Ables, S.T., Fraser, B.J., Waters, C.L. and Neudegg, D.A. (1998). Monitoring cusp/cleft topology using Pc5 ULF waves. *Geophysical Research Letters*, 9: 1507–1510.

Alfvén, H. (1950). *Cosmical Electrodynamics*. Oxford University Press, New York.

Ansari, I.A. and Fraser, B.J. (1986). A multistation study of low latitude Pc3 geomagnetic pulsations. *Planetary and Space Science*, 34: 519–536.

Bol'shakova, O.V. and Troitskaya, V.A. (1984). The relation of the high-latitude maximum of Pc3 intensity to the dayside cusp. *Geomagnetism and Aeronomy*, 24: 633–635.

Fukunishi, H. and Lanzerotti, L.J. (1974a). ULF pulsation evidence of the plasmapause, 1, Spectral studies of Pc3 and Pc4 pulsations near L = 4. *Journal of Geophysical Research*, 79: 142–158.

Fukunishi, H. and Lanzerotti, L.J. (1974b). ULF pulsation evidence of the plasmapause, 2, Polarization studies of Pc3 amd Pc4 pulsations near L = 4 and at a latitude network in the conjugate region. *Journal of Geophysical Research*, 79: 4632–4647.

Greenstadt, E.W., Mellot, M.N., McPherron, R.L., Russell, C.T., Singer, H.J. and Krecht, D.J. (1983). Transfer of pulsation related wave activity across the magnetopause: observations of corresponding spectra by ISEE-1 and ISEE-2. *Geophysical Research Letters*, 10: 659–662.

Gustafsson, G., Papitashvili, N.E. and Papitashvili, V.O. (1992). A revised corrected geomagnetic coordinate system for epochs 1985 and 1990. *Journal of Atmospheric and Terrestrial Physics*, 54: 1609–1631.

Howard, T.A and Menk, F.W. (2001). Propagation of 10–50 mHz ULF waves with high spatial coherence at high latitudes. *Geophysical Research Letters*, 28: 231–234.

Kato, K., Sato, N. and Tonegawa, Y. (1994). Estimation of geomagnetic conjugate points using Pc5 geomagnetic pulsations. *Journal of Geomagnetism and Geoelectricity*, 46: 363–372.

Lanzerotti, L.J., Hasegawa, A. and Tartaglia, N.A. (1972). Morphology and interpretation of magnetospheric plasma waves at conjugate points during December solstice. *Journal of Geophysical Research*, 77: 6731–6745.

Mathie, R.A., Mann, I.R., Menk, F.W. and Orr, D. (1999). Pc5 ULF pulsations associated with waveguide modes observed with the IMAGE magnetometer array. *Journal of Geophysical Research*, 104: 7025–7036.

Menk, F.W., Fraser, B.J., Waters, C.L., Ziessolleck, C.W.S., Feng, Q., Lee, S.H. and McNabb, P.W. (1994). Ground measurements of low latitude magnetospheric field line resonances. *AGU Monagraph*, 81: 299–310.

Olson, J.V. and Szuberla, C.A.L. (1997). A study of Pc3 coherence at cusp latitudes. *Journal of Geophysical Research*, 102: 11375–11383.

Samson, J.C. and Rostoker, G. (1972). Latitude-dependent characteristics of high-latitude Pc4 and Pc5 micropulsations. *Journal of Geophysical Research*, 77: 6133–6144.

Sugiura, M. and Wilson, C.R. (1964). Oscillations of the geomagnetic field lines and associated magnetic perturbations at conjugate points. *Journal of Geophysical Research*, 69: 1211–1216.

Tonegawa, Y. and Fukunishi, H. (1984). Harmonic structure of Pc3-5 magnetic pulsations observed at the Syowa-Hasafell conjugate pair. *Journal of Geophysical Research*, 89: 6737–6748.

Tsyganenko, N.A. (1989). A magnetospheric magnetic field model with a warped tail current sheet. *Planetary and Space Science*, 37: 5–20.

Waters, C.L., Menk, F.W. and Fraser, B.J. (1991b). The resonance structure of Pc3 geomagnetic pulsations. *Geophysical Research Letters*, 18: 2293–2296.

Waters, C.L., Samson, J.C. and Donovan, E.F. (1995). The temporal variation of the frequency of high latitude field line resonances. *Journal of Geophysical Research*, 100: 7987–7996.

Ziesolleck, C.W.S. and McDiarmid, D.R. (1995). Statistical survey of auroral latitude Pc5 spectral and polarization characteristics. *Journal of Geophysical Research,* 100: 19299–19312.

Ziesolleck, C.W.S., McDiarmid, D.R. and Feng, Q. (1997). A comparison between the Pc3-4 pulsations observed by the GOES 7 and the CANOPUS magnetometer array. *Journal of Geophysical Research,* 102: 4893–4909.

26. OUTER MAGNETOSPHERE TOPOLOGY AND PC5 ULF WAVES AT HIGH-LATITUDES

S.T. Ables[1], B.J. Fraser[1], F.W. Menk[1], D.A. Neudegg[1,4], R.J. Morris[2] and Liu Ruiyuan[3]

[1] Department of Physics
Cooperative Research Centre for Satellite Systems
University of Newcastle
NSW 2308 Australia
(email: phbjf@cc.newcastle.edu.au)

[2] Atmospheric and Space Physics
Australian Antarctic Division
Kingston Tasmania 7050 Australia

[3] Polar Research Institute of China
Shanghai 200129
China

[4] now at:
Space Physics Centre
University of Leicester UK

Abstract

Induction magnetometer data recorded at three closely spaced sites (110 km) in Antarctica (mlat -75°) have been examined for ionospheric signatures of the outer magnetosphere and cusp/cleft regions. Cross-phase analysis of the 1–10 mHz band, using pure-state filtering techniques reveal diurnally varying field line resonances embedded in the spectra, while interstation phase lag measurements indicate azimuthal propagation of waves away from local magnetic noon. Using the T96 (Tsyganenko, 1995) external field model, cross-phase measurements are put in the context of diurnally changing field line topology, due to compression at the sub-solar region and stretching along the dawn and dusk flanks. This suggests field line stretching as an alternative or additional explanation to the Kelvin-Helmholtz instability for azimuthal wave propagation. A 'wave cusp meridian' identified in the propagation data shows a variable location about local magnetic noon which is under the control of IMF By and shows the same well-known behaviour as the cusp to these variations.

26.1 Introduction

Signatures of the low altitude cusp have become an essential tool in our understanding of the dynamic interaction between the solar wind and

the Earth's magnetosphere. Many instruments located at cusp latitudes have been employed to identify such signatures. Intercalibration between the particle, optical and radar cusps is an ongoing project with many successes as well as controversies (e.g. Newell, 1994). The role of ground-based magnetometers in this quest is often secondary due to low spatial resolution and limited coverage at suitable latitudes. In recent years large magnetometer arrays such as MACCS in Canada and MAGIC in Greenland, coupled with interstation spectral analysis techniques (e.g. Waters *et al.*, 1995) have improved the utility of these instruments.

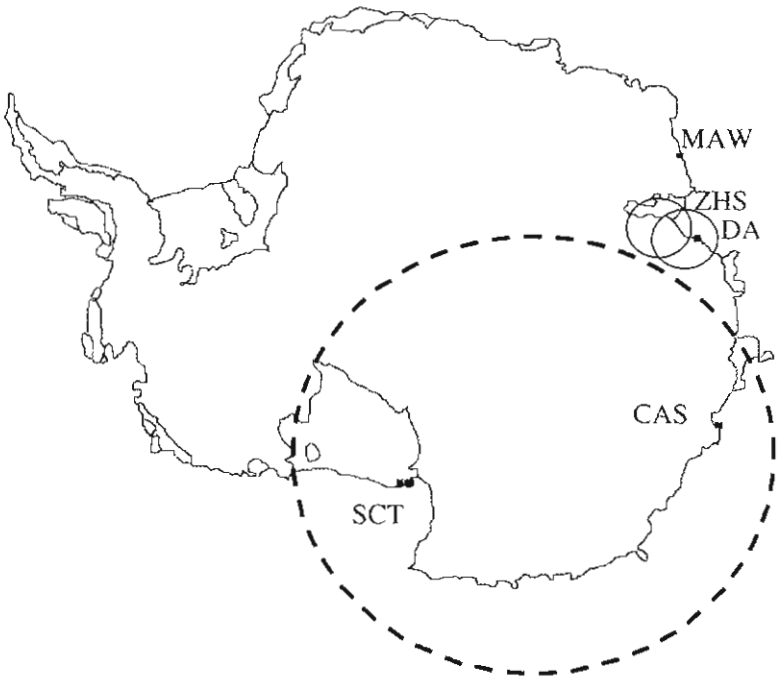


Figure 1. *Location of Newcastle University/Australian Antarctic Division induction magnetometers on the Antarctic continent. The dotted line shows the average geomagnetic latitude of the particle cusp. The 200 km field-of-view circles are shown around Davis and Zhong Shan magnetometers.*

Antarctica, with its harsh environment and geographical isolation has not been as heavily exploited as a platform for such arrays, although the deployment of Automatic Geophysical Observatories (AGO's) across the Antarctic plateau is providing new opportunities for southern hemisphere high-latitude studies. The University of Newcastle in conjunction with the Australian Antarctic Division has maintained an array of induction magnetometers along the eastern Antarctic coastline for many years (see Figure 1). In 1992 two temporary sites were added at cusp latitudes

near the permanent Australian station Davis. A successful campaign was performed with this triangle of stations to identify the source and propagation characteristics of Pc1-2 (100–500 mHz) geomagnetic pulsations (Neudegg *et al.*, 1995). This data set was later re-analysed to characterise Pc5 (1–10 mHz) pulsations with the aim of identifying the 'wave cusp' (Ables *et al.*, 1998). Although the data set was limited to two weeks, two key results were obtained:

(i) Daily reversals of interstation phase lags were readily identified in magnetically east-west spaced stations around the time of local magnetic noon.

(ii) Field line resonances were detected on most days indicating that the stations lay under closed field lines. The occasional disappearance of these resonances under solar wind conditions of large negative IMF B_z further indicated these were the last closed field lines in the dayside magnetosphere.

In 1996 a more permanent magnetometer installation was undertaken at the Chinese base of Zhong Shan, ~ 110 km geomagnetic west of Davis (see Figure 1), and further examinations of these 'wave cusp' diagnostics have been pursued. This paper is a preliminary study of these data.

26.2 The data

The geomagnetic coordinates of Davis and Zhong Shan are 74.6°S, 102.37°E and 74.7°S, 98.63°E (PGM-88; Baker and Wing, 1989) respectively, and these stations are ideally located for cusp studies under quiet to moderate geomagnetic activity. Magnetic variations were recorded at each station using two orthogonal induction coils. The outputs from the coils were filtered and preamplified to give the time derivatives of the H and D components (X, Y) of the local magnetic field. The sensitivity of the combined detection and logging system was 0.2 nT V^{-1} at 10 mHz, with the frequency response falling at a rate of 6 dB per octave down to 1 mHz (Fraser *et al.*, 1991). This had the effect of spectrally flattening the background $f^{-2.6}$ power law spectrum (Olson, 1986) of the geomagnetic pulsations. The phase response over this band was essentially flat and timing at each station was provided by GPS measurements. The analogue output from the magnetometers was digitised using an ADAS data processing unit with a sample rate of 0.5 s^{-1} at each station (Symons, 1996).

26.3 Data analysis

All data were low pass (anti-alias) filtered and resampled at a rate of 1/20 s^{-1}, providing a Nyquist frequency of 25 mHz, suitable for analysis of Pc5 band (1–10 mHz) pulsations. Dynamic spectrograms were generated using a 256 point (85 minute) Fast Fourier Transform (FFT) stepped at 10 minute intervals. In order to place the data in context with respect to the general frequency structure of Pc5 waves in the cusp region, we have shown in Figure 2 dynamic spectra for eight consecutive days over 1–8 April 1996, which show varying geomagnetic

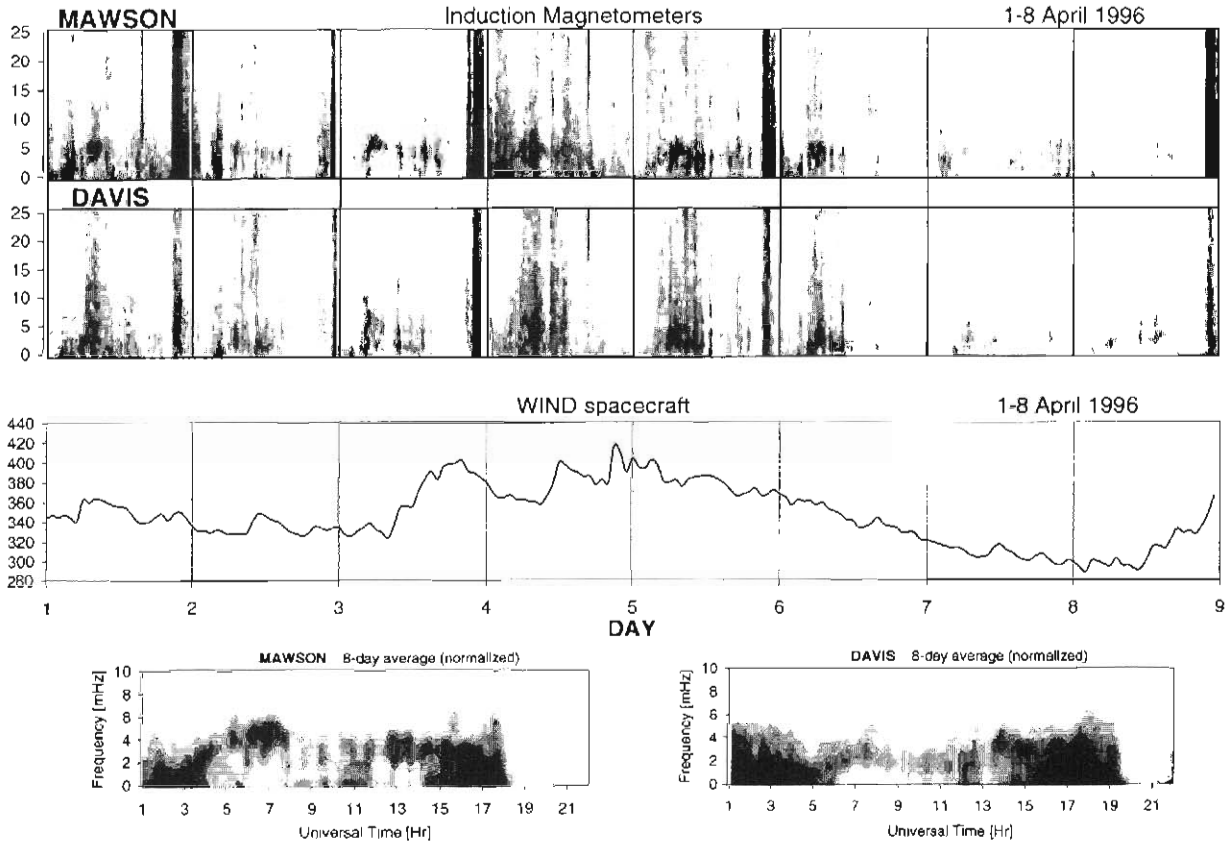


Figure 2. Cross-power (X, Y) dynamic spectral data from Mawson and Davis over eight consecutive days. Also plotted is the WIND spacecraft solar wind velocity over these days. The bottom panels are averaged dynamic spectral plots for the eight days at Mawson and Davis emphasising the 'Pc5 arch'. Local noon is 0940 UT at Davis and 1020 UT at Mawson.

activity. Only two stations, Davis and Mawson are shown. Zhong Shan, because of its close proximity, shows spectra essentially identical to Davis. Additional data were available from Mawson which is located ~ 600 km geomagnetically north-west of Davis at an auroral zone latitude (70.5°S, 92.1°E; PGM-88). Figure 2 shows the X component power at Davis and Mawson over eight consecutive days where geomagnetic activity is slowly decreasing along with the solar wind velocity measured by WIND satellite. The plots in the top two rows of panels show dynamic power spectra over a full 24 hour UT day. These include both coherent and incoherent power and the spectra over 1, 4 and 6 April are similar to those presented by Engebretson *et al.* (1995) which were dominated by Pi1 noise. In general the integrated magnetometer power (1–25 mHz) at both stations roughly follows the solar wind velocity variations measured by WIND, with increased power corresponding to higher solar wind velocity. On 4 April, where ΣK_p over the day exceeded 20, there is no discernable frequency structure in the spectra other than the most intense activity seen between 0500–1300 UT around or prior to local magnetic noon (~ 0940 UT).

On 1, 4 and 5 April substorm activity occurs on the nightside after about 2000 UT. Over 2, 3 and 6 April, with moderate to low geomagnetic activity the daytime Pi1 noise reduces in intensity, revealing the 'Pc5-arch' spectra originally defined by McHarg and Olson (1991). The frequency structure of the arch was originally attributed by McHarg and Olson (1991) to the passage of the cusp overhead in local time. However, Waters *et al.* (1995) using CANOPUS magnetometer chain data identified the arch as a high-latitude signature of field line resonance. The arch straddles local noon and is clearly seen at Mawson between 0300–1100 UT and possibly extending to 1500 UT, with a peak frequency in the range 3–5 mHz in the 8-day averaged spectra shown in the bottom panels of Figure 2. The arch is more difficult to identify at Davis but is apparent in the 8-day averaged plot with a frequency of 2–4 mHz. The decrease in arch frequency with increasing latitude is in agreement with the properties of a field line resonance.

The data presented in Figure 2 indicates how the Pi noise obliterates the field line resonance signature at high-latitudes near the cusp on geomagnetically active days. If the Pi signals are incoherent and the 'Pc5 arch' is coherent and shows stable polarisation, then it should be possible to use sophisticated analysis techniques to further study the Pc5 waves. In order to achieve this, pure-state filtering techniques (Samson, 1983) were used to detect polarised components in the pulsations.

As a test of our cross-phase analysis we first determined the horizontal polarisation of Pc5 waves recorded at Davis. We confirmed the observations made by Samson (1972) and more recently by Olson and Fraser (1994), of a diurnal variation in ellipticity from right-hand before magnetic noon to left-hand/linear after magnetic noon. An example of this is seen in the display of eight days of data over 1–8 April 1996 in

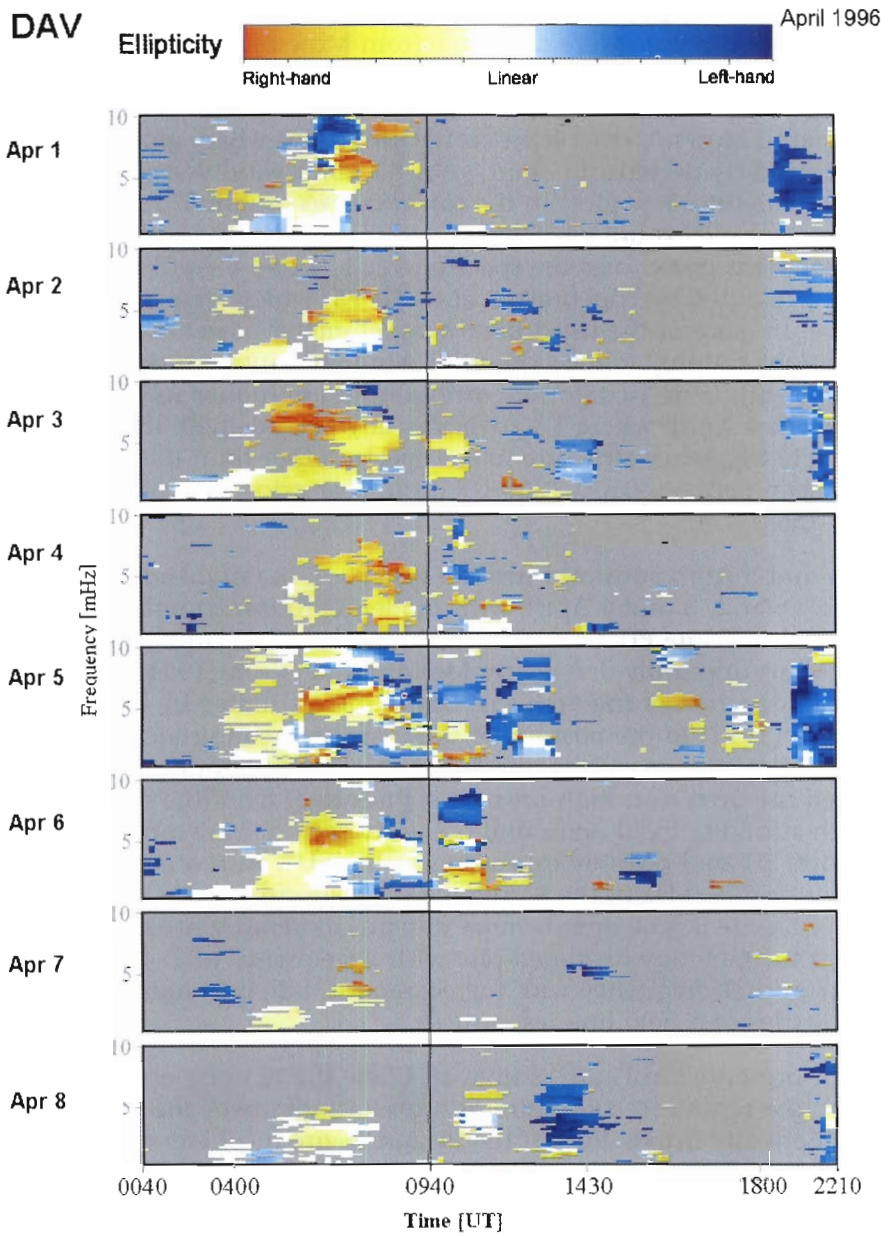


Figure 3. *Pure state polarisation properties at Davis over 1–8 April 1996. Data are plotted only where polarised power is greater than 70% of the total power. Right-hand polarisation dominates pre-noon and left-hand post noon.*

Figure 3, where the pure-state polarisation properties in the plane of the Earth's surface (X–Y) at Davis are plotted as dynamic ellipticity spectra. Data with polarised power < 70% are not plotted here. Right-hand polarisation, defined along the direction of the Earth's field, is indicated by negative phase and dominates in the pre-noon period (white to orange). The post-noon period generally shows the opposite left-hand polarisation (blue to purple) but the coherent signals here are patchy and less distinct than the morning signals. Significantly, the polarisation of Pc5 waves recorded at the auroral oval station, Mawson, during the same period showed a reversal of this trend, in agreement with Samson (1972). The theoretical interpretations made by Chen and Hasegawa (1974), together with Southwood (1974), described above, point to the presence of field line resonances driven by azimuthally propagating boundary waves.

26.4 Azimuthal propagation

The group of eight days was also analysed with respect to interstation phase. The results are shown in the daily phase lag spectrograms in Figure 4. Here the cross-phase between the X component signals at Davis and Zhong Shan are plotted. Since these stations are at similar latitudes the cross-phase lags indicate azimuthal propagation of Pc5 waves in the 1–10 mHz band. Phase lags between geomagnetic east-west separated cusp latitude stations such as Davis and Zhong Shan can be interpreted as propagation of waves at the magnetosphere mapped along field lines to the ground via the Alfvén/fast mode (e.g. Olson and Rostoker, 1978). A change in the sign of the Pc5 phase is seen in Figure 4 to occur around 0940 UT. Here, a negative (positive) phase indicating westward (eastward) azimuthal propagation is observed in the local morning (afternoon). The sign of the phase lags and their reversal at a time close to local magnetic noon (0940 UT) indicate propagation away from the subsolar point and toward the magnetotail along the dawn and dusk flanks of the magnetosphere. This would be the case if waves were formed, for instance, by the Kelvin-Helmholtz instability in the boundary layer.

Accurate daily estimates of the time of noon reversals can be made using the criterion that the slope of phase versus frequency changes sign for a range of frequencies centred around 5 mHz. Although the phase reversal is seen, on average around Davis local magnetic noon at about 0940 UT, it is possible to track the precise reversal on a day-to-day basis. This is shown by the crosses in Figure 4 and this follows the phase-

that the fundamental field line resonance frequency of the arch is 2–4 mHz at Davis. Figure 4 indicates phase reversals at these and higher frequencies. The higher frequency reversals may be associated with field line resonance harmonics which have been identified in Davis data by Ables *et al.* (1998), with the second harmonic at 4–7 mHz.

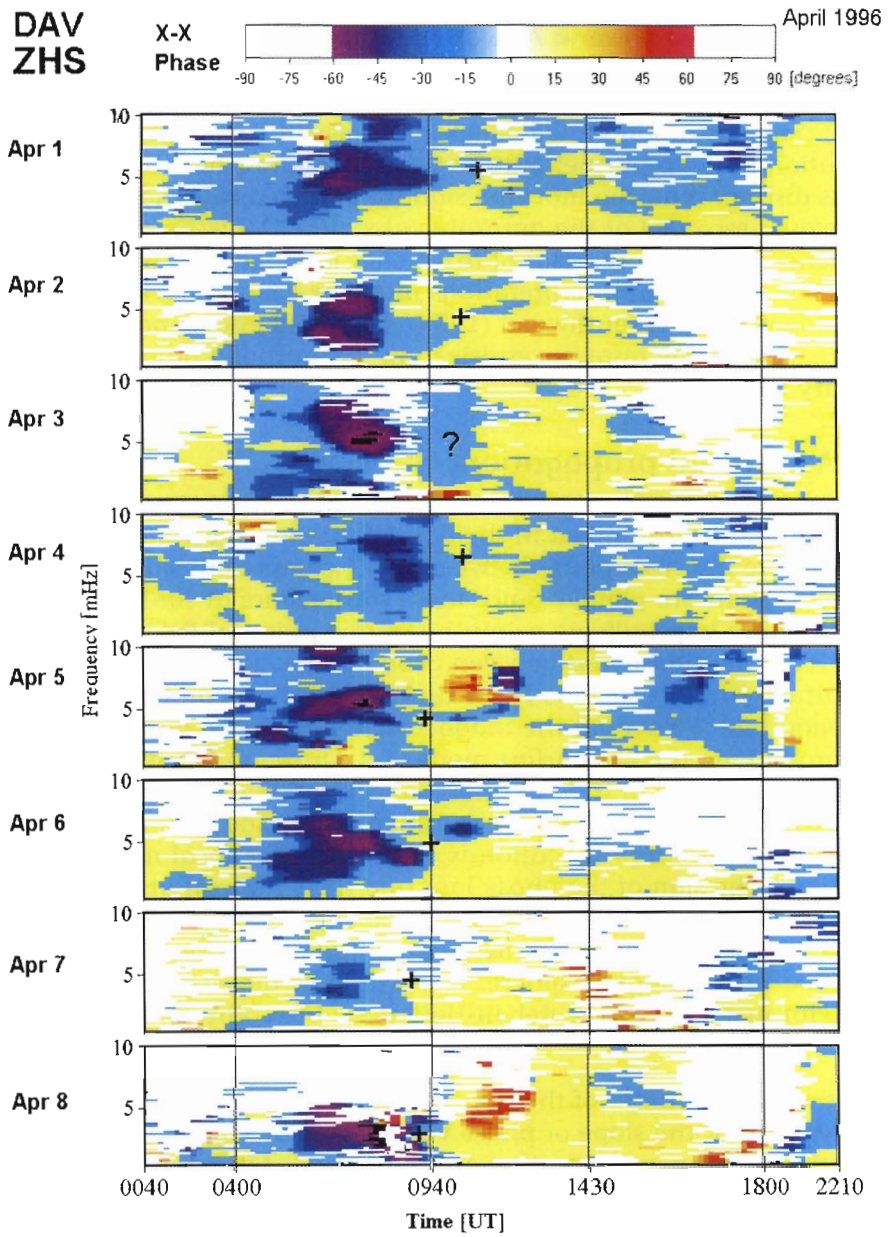


Figure 4. *Pure state daily cross-phase difference dynamic spectra for the X component signals between Davis and Zhong Shan over 1–8 April 1996. Phase differences are in the range ~ 60° and only this is plotted. The plus signs seen in the cross-phase spectra near local noon (0940 UT) identify the passage of the cusp meridian. Vertical lines indicate phase reversals. Right-hand polarisation dominates pre-noon and left-hand post-noon.*

Using the PGM-88 geomagnetic coordinate system, local magnetic noon at Davis should occur at 0955 UT, and at Zhong Shan at 1010 UT. Comparison with the estimates described above indicate that the wave cusp meridian occurs ~ 20 minutes earlier than these times. We follow the argument of Olson and Rostoker (1978) in interpreting this as an aberration of the solar wind direction ~ 5° away from the sub-solar point due to the Earth's orbital motion. It is also of interest to look for other patterns in phase reversals related to local time. From Figure 4 it can be seen that on most days a corresponding reversal occurs ~ 2 hours before local magnetic midnight around 1950 UT. This is possibly due to propagation of Pi2 pulsations away from the substorm generation region. In this case the sign of the phase lag indicates sunward propagation. These azimuthal propagation properties are illustrated in Figure 5 which shows a plot of the geomagnetic field line topology in the equatorial plane in GSM coordinates using the T-96 model (Tsyganenko, 1995) under conditions prevalent on 5 April 1996 with D_{st} = 0, solar wind dynamic pressure = 2 nPa, and IMF (B_y; B_z) = 0. The T96 model has a realistic magnetopause imposed and was incorporated with the IGRF internal field. The field lines shown thread Davis with the grey curve indicating the equatorial magnetopause and the arrows showing the direction of azimuthal propagation associated with the Pc5 or Pi2 phase reversals deduced from Figure 4. These reversals are reasonably

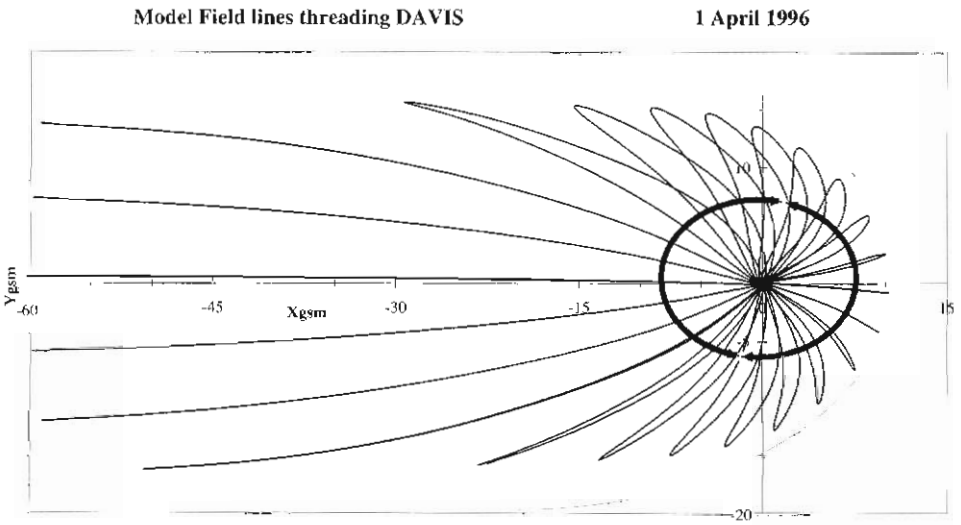


Figure 5. *Field lines threading Davis on 8 April 1996 completed from the Tsyganenko T-95 model. The equatorial magnetopause is also shown. The arrows indicate the direction of phase lag from Figure 4.*

consistent across all eight days shown in Figure 4 indicated by the four vertical lines.

Phase versus frequency relations in the pre-noon sector (0330–0940 UT) cannot be explained by propagation alone, as the phases show a minima around 3 mHz rather than a linear relation with frequency. We resolve this by invoking field line resonance.

26.5 Field line resonance

To identify field line resonances in our data we first assume there is an ionospheric signature within the common field-of-view of two magnetometers. The cross-power of horizontal magnetic variations detected from each magnetometer should then show a spectral peak at the resonant frequency. Further, if the resonance is sufficiently broad (i.e. slow spatial variation compared to station spacing, low Q) then the interstation cross-phase should likewise peak at the resonant frequency. Further details of this technique are given in Waters *et al.* (1991) and Chi and Russell (1998).

Using the WKB approximation for estimating the resonant frequency and assuming that the plasma density is reasonably constant in the outer plasmatrough, Ables *et al.* (1998) postulated that variations in frequency are due to changes in field line length, as shown in Figure 5. This frequency shift determines the cross-phase at the resonant frequency common to both the Davis and Zhong Shan magnetometers, which have overlapping fields-of-view (Figure 1). Near resonance the phase is a linear function of frequency (Waters *et al.*, 1995). Therefore the cross-phase between closely spaced stations will be proportional to the fractional frequency and the change in field line lengths (see Equation 4 in Ables *et al.*, 1998).

We can therefore have an alternate explanation to the Kelvin-Helmholtz instability for the phase reversals seen at local magnetic noon (LMN). Before LMN field lines threading Davis are longer than those threading Zhong Shan (see Figure 5), therefore from the above argument we expect a negative phase lag. After LMN the situation is reversed. In order to confirm this argument more quantitative analysis is required to separate the effects of propagation and resonance.

26.6 Solar wind control

The solar wind conditions observed by WIND over the eight day period of the study are shown in Figure 6 and were examined in conjunction with the corresponding phase lag dynamic spectrograms illustrated in Figure 4. A number of interesting observations can be obtained by comparing the two data sets. The maximum negative IMF B_z occurred on 4 April and it is important to note that field line resonances were still present in the pre-noon sector on this day, suggesting the stations remained under closed field lines. The IMF B_y trend in Figure 6 is from negative to positive over the eight days, this is accompanied by

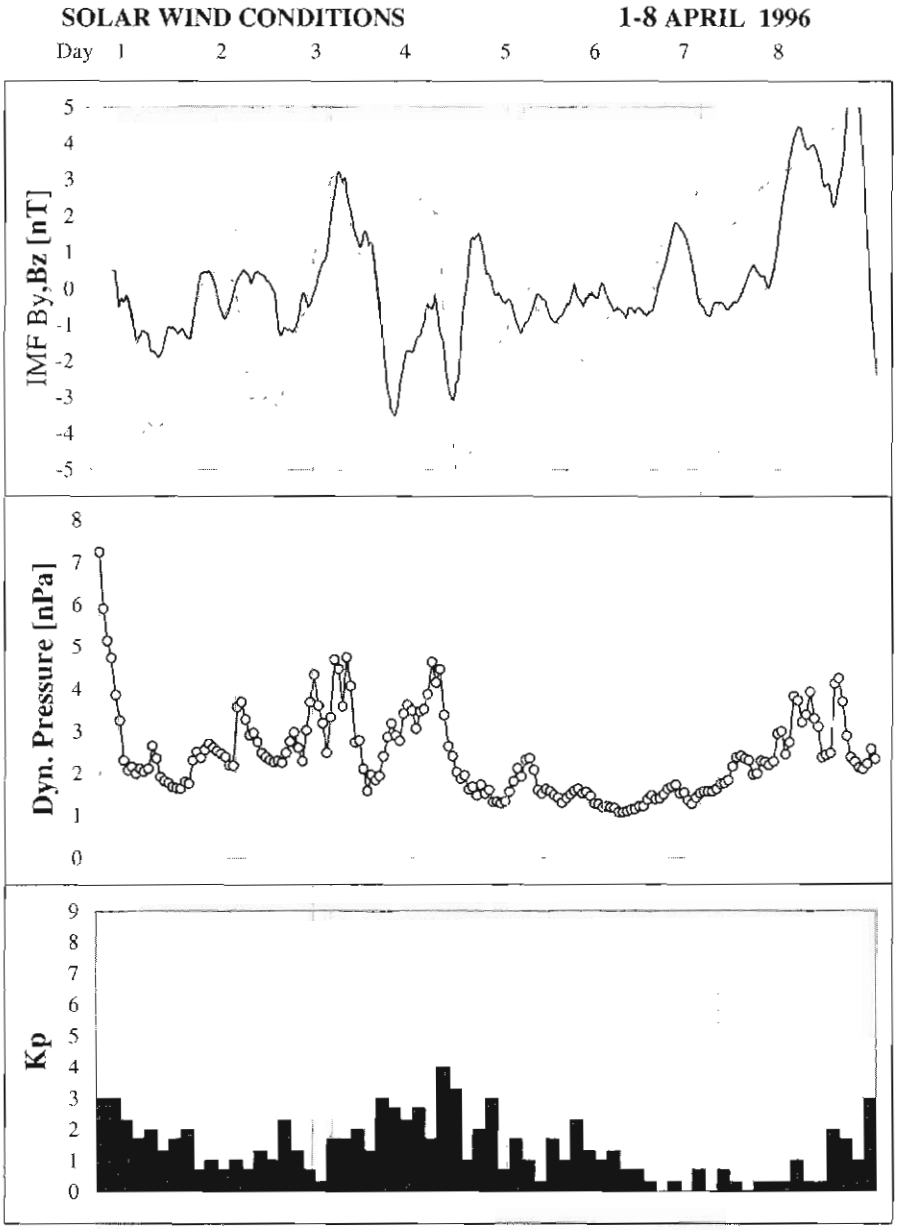


Figure 6. *IMF magnetic field data from WIND key parameters. Solar wind dynamic pressure is adapted from WIND-SWE key parameters. (R. Lepping, PI). Vertical dotted lines indicate intervals between 0600 and 1030 UT on each day over 1–8 April 1996. Kp for these days is plotted in the bottom panel.*

a corresponding trend in the time of the 'wave cusp meridian' passage, indicated by the crosses in Figure 4, from late to early with respect to the average. A similar response in the location of the dayside cusp in the pre-noon sector for IMF B_y negative and in the post noon for IMF B_y positive is well known (e.g. Cowley, 1981). Peaks in solar wind dynamic pressure on 3 and 8 April 1996 result in greater field line stretching along the flanks of the magnetopause. We note that increased gradients in field line lengths result in deeper pre-noon phase minima, as indicated in Figure 4 by the presence of grey patches with a phase lag < -90°.

Days 5 and 6 April 1996 have similar and reasonably quiet solar wind conditions and exhibit a similarity between the phase lag dynamic spectrograms in Figure 4. It is apparent that solar wind and IMF control plays an important role in determining the detail of the spectral and phase structure of Pc5 field line resonances in the outer plasma trough.

26.7 Summary

This paper shows how field line resonance may be used to indicate properties of the magnetosphere using inexpensive induction magnetometers. It is important to realise that the wave features illustrated in Figures 4 and 5 are a consequence of using two channel pure-state analysis techniques, to eliminate background incoherent Pi1 'noise' of the type dominating the spectra shown for active days in Figure 2. In particular, with two stations closely spaced in azimuth by ~ 110 km and on closed field lines just inside the magnetopause, it has been possible to monitor the geomagnetic field topology through the analysis of Pc5 wave phase and azimuthal propagation. This has led to the identification of a 'wave cusp meridian' which shows a variable location about local magnetic noon under the control of IMF B_y. Another important result is the suggestion that field line stretching either side of magnetic noon may contribute to azimuthal phase shifts and provide an alternative or additional explanation to the Kelvin-Helmholtz instability.

The results reported are part of the ULF wave program at the University of Newcastle to further understand processes in the high-latitude ionosphere-magnetosphere system and to develop diagnostic techniques of use in space weather modelling.

26.8 Acknowledgments

The continued maintenance of equipment, and collection of data from the Australian Antarctic stations has been made by the Atmospheric and Space Physics expeditioners and logistical support of the Australian Antarctic Division. Appreciation is expressed to the CHINARE expeditioners who diligently maintained the equipment at Zhong Shan. The research was supported by grants from the

Australian Research Council, Antarctic Science Advisory Committee and the University of Newcastle.

References

Ables, S., Fraser, B.J., Menk, F.W. and Morris, R.J. (1998). Monitoring cusp/cleft topology using Pc5 ULF waves. *Geophysical Research Letters*, 25: 1507–1510.

Baker, K.B. and Wing, S. (1989). A new coordinate system for conjugate studies at high latitudes. *Journal of Geophysical Research*, 94: 9139–9143.

Chen, L. and Hasegawa, A. (1974). A theory of long-period magnetic pulsations. 1. Steady state excitation of field line resonances. *Journal of Geophysical Research*, 79: 1024–1032.

Chi, P. and Russell, C.T. (1998). An interpretation of cross-phase spectrum of geomagnetic pulsations by field line resonance theories. *Geophysical Research Letters*, 25: 4445–4448.

Cowley, S.W.H. (1981). Magnetospheric asymmetries associated with the y-component of the IMF. *Planetary and Space Science*, 29: 79–96.

Engebretson, M.J., Hughes, W.J., Alford, J.L., Zesta, E., Cahill Jr, L.J., Arnoldy, R.L. and Reeves, G.D. (1995). Magnetometer array for cusp and cleft studies: observations of the spatial extent of broadband ULF magnetic pulsations at cusp/cleft latitudes. *Journal of Geophysical Research*, 100: 19371–19386.

Fraser, B.J., McNabb, P.W., Menk, F.W. and Waters, C.L. (1991). A personal computer induction magnetometer system for recording geomagnetic pulsations. In: Burns, G. (Ed.). *ANARE Research Notes*, 80. Australian Antarctic Division. Pp. 83–92.

McHarg, M.G. and Olson, J.V. (1992). Correlated optical and magnetic observations of the winter cusp-boundary layers system. *Geophysical Research Letters*, 19: 817–820.

Neudegg, D.A., Fraser, B.J., Menk, F.W., Hansen, H.J., Burns, G.B., Morris, R.J. and Underwood, M.J. (1995). Sources and velocities of Pc1-2 ULF waves at high latitudes. *Geophysical Research Letters*, 22: 2965–2968.

Newell, P.T. (1995). Do the dayside cusps blink?, U.S. National Report to IUGG, 1991–1994. Reviews of Geophysics 33, 665–668.

Olson, J.V. (1986). ULF signatures of the polar cusp. *Journal of Geophysical Research*, 91: 10055–10062.

Olson, J.V. and Rostoker, G. (1978). Longitudinal phase variations of Pc 4-5 micropulsations. *Journal of Geophysical Research*, 83: 2481–2488.

Samson, J.C. (1972). Three dimensional polarisation characteristics of high-latitude Pc5 geomagnetic micropulsations. *Journal of Geophysical Research*, 77: 6145–6160.

Southwood, D.J. (1974). Some features of field line resonance in the magnetosphere. *Planetary and Space Science*, 22: 483–491.

Symons, L.P. (1996). An analogue data acquisition system for ANARE stations. In: Morris, R.J. (Ed.). *ANARE Research Notes 95*. Australian Antarctic Division. Pp. 168–176.

Tsyganenko, N.A. (1995). Modeling the Earth's magnetospheric magnetic field confined within a realistic magnetopause. *Journal of Geophysical Research*, 100: 5599–5612.

Waters, C.L., Menk, F.W. and Fraser, B.J. (1991). The resonant structure of low latitude Pc3 geomagnetic pulsations. *Geophysical Research Letters*, 18: 2293–2296.

Waters, C.L., Samson, J.C. and Donovan, E.F. (1995). Temporal variation of the frequency of high latitude field line resonances. *Journal of Geophysical Research*, 100: 7987–7996.

27. HF DOPPLER OSCILLATIONS DUE TO MIXED ULF WAVE MODES

C.L. Waters, M. Sciffer, I.S. Dunlop and F.W. Menk

Department of Physics
Cooperative Research Centre for Satellite Systems
The University of Newcastle
Callaghan NSW 2308 Australia
(email: physpuls8@cc.newcastle.edu.au)

Abstract

HF signals that propagate via the ionosphere may exhibit Doppler shifts due to a number of processes that give rise to a time dependent refractive index. This paper examines the role of ULF (1–100 mHz) wave energy, incident from the magnetosphere, in causing Doppler shifts in HF signals. We describe a model of ULF wave propagation through the ionosphere that allows the ULF wave mode mix to be specified. It is found that the shear Alfvén mode wave gives the larger Doppler velocities of order 2 m s^{-1} in the F-region, caused mostly by a mechanism that involves the divergence of the electron velocity from the passage of the ULF wave. Pure fast mode perturbations of similar amplitude gives Doppler velocities around ten times smaller. In both cases the role of vertical bulk motion of the electrons is small. An experiment that recorded data from a Doppler sounder and magnetometer is described and the amplitude-ratio and cross-phase of data recorded by the north-south magnetometer sensor and the Doppler sounder are presented. The amplitude-ratio data show a distinctive maximum and the cross-phase a large phase shift around a ULF field line resonance. The paper presents the relevant theory and modelling of this feature and illustrates the dependence on the horizontal spatial structure of the ULF energy.

27.1 Introduction

The role of the ionosphere in determining propagation properties of HF (3–30 MHz) signals has been studied since the advent of radio. At appropriate frequencies, the ionospheric plasma causes HF signals to be continually refracted. This makes long distance communications possible yet very dependent on ionospheric parameters. The ionosphere is birefringent due to the imposed geomagnetic field which also supports ultra-low frequency (ULF, 1–100 mHz) wave energy incident from the magnetosphere. Near Earth space is a magnetised plasma environment where ULF energy propagates as magnetohydrodynamic (MHD) waves which can be measured using sensitive magnetometers. When ULF waves interact with the ionosphere, they cause frequency shifts in any HF signals present.

The relationship between ULF waves recorded by ground-based magnetometers and Doppler variations in ionosphere propagating HF waves was first noticed by Harang (1939). Since then, the effect has been extensively examined experimentally (Watermann, 1987 and references therein). A mathematical description was formulated by Rishbeth and Garriott (1964) who proposed two mechanisms for the observed Doppler effects. The first involved a polarisation electric field generated in the E-region, influencing the F-region as an E×B drift. The second mechanism describes bulk motion of the F-region due to the MHD wave. Jacobs and Watanabe (1966) improved the model by including changes in the refractive index due to variations in the ionospheric electron distribution. A more complete and rigorous theory has been developed in a series of papers by Poole and Sutcliffe (1987), Poole *et al.* (1988) and Sutcliffe and Poole (1989, 1990) which we will denote as the SP model. This model identifies three mechanisms that might alter the refractive index and lead to Doppler shifts in HF signals. In this paper we use the latest modifications of the SP model described by Sutcliffe and Poole (1989, 1990). A number of issues that involve the electron collision terms, complex refractive index and electron velocity terms are discussed.

The HF signal propagation path, s, in the ionosphere is determined by the signal frequency, f, and the medium refractive index, n. Any time variation in the refractive index will also vary the propagation path, giving rise to a Doppler frequency shift, Δf, of the HF signal. For φ, the angle between the direction of HF signal energy transport and the wave normal, the frequency shift is given by

$$\Delta f = -\frac{f}{c}\frac{d}{dt}\int_s n(s)\cos\varphi ds \qquad (1)$$

In this paper, Doppler shifts in HF signals arising from changes in refractive index associated with ULF (1–100 mHz) waves incident from the magnetosphere are discussed. The main result is the extension of previous theoretical work to include effects from a wider range of incident ULF wave modes. The paper also highlights how the combination of HF and ULF data can provide considerable insight into ionospheric processes and the effects of incident ULF energy on HF signal propagation.

27.2 ULF waves through the ionosphere

ULF wave energy propagates through the magnetised plasma environment of the Earth's magnetosphere where the ionosphere represents the inner boundary. The energy source for these waves can be traced ultimately to solar activity, while closer to Earth, the bow shock and magnetopause are ULF wave generation regions. ULF energy in the cold plasma of the magnetosphere exists in two wave modes, known as the fast MHD and shear Alfvén wave modes. The presently accepted, although debated, scenario is for fast mode energy, that can propagate across the magnetic field, to mode convert and excite the shear Alfvén mode as a resonant oscillation between conjugate ionospheres (Chen and

Hasegawa, 1974; Southwood, 1974; Yumoto *et al.*, 1985). An important property of the shear Alfvén mode is that the wave energy is magnetic field line guided even if the propagation vector, k, has a component across the field. The shear Alfvén mode energy can reflect at conjugate ionospheres to form field line resonances (FLR's) where the frequency depends on the plasma density and magnetic field topology and hence latitude. This means that in the 1–100 mHz band for a detector at a given latitude, FLR's occupy a small section of the spectrum so that most of the energy is in the fast mode or some mixture of the two modes.

The procedure for determining HF Doppler shifts due to ULF wave energy in the ionosphere requires a knowledge of the ULF wave fields as a function of height from the ground up to the HF reflection point. The SP papers used the model of Hughes (1974), and Hughes and Southwood (1976) to determine the ULF wave fields through the ionosphere. This model is formulated as an initial value problem but has difficulties with numerical swamping (e.g. Pitteway, 1965). The SP papers also limited the application to shear Alfvén mode waves. A more recent and flexible formulation of the problem of ULF wave propagation through the ionosphere has been given by Zhang and Cole (1994, 1995). This recasts the equations as a boundary value problem and allows greater flexibility for tailoring the boundary conditions, particularly the incident ULF wave properties. A variant of this scheme is described and used in this paper.

Assuming the medium changes only in the vertical direction (z), the relevant Maxwell equations are

$$\nabla \times \mathbf{E} = -i\omega\mathbf{B} \tag{2}$$

$$\nabla \times \mathbf{B} = \mu_0\mathbf{J} + i\omega\mu_0\varepsilon_0\mathbf{E} \tag{3}$$

where

$$\mathbf{J} = \overrightarrow{\sigma}(z)\cdot\mathbf{E}\ . \tag{4}$$

The formulation of the problem for any coordinate system is achieved through the terms in the conductivity tensor, σ. By taking the curl of Equation (2) and substituting in Equation (3), Zhang and Cole (1994) derived equations for the horizontal electric field quantities in terms of two second order differential equations with eight rather involved coefficients. Keeping Equations (2) and (3) separate and formulating the problem as a system of first order equations appeared to be a much simpler route. For the permittivity tensor related to the conductivity by

$$\overrightarrow{\varepsilon} = \overrightarrow{I} - \frac{i}{\omega\varepsilon_0}\overrightarrow{\sigma} \tag{5}$$

and a vertical ambient magnetic field, the relevant equations for the ULF perturbation electric and magnetic fields are

$$\frac{db_x}{dz} = E_x\left[\frac{i}{c}\frac{k_x k_y}{k_0} + \varepsilon_2\frac{k_0}{c}\right] + E_y\left[\frac{ik_0}{c}\varepsilon_1 - \frac{ik_x^2}{c}\right] \tag{6}$$

$$\frac{db_y}{dz} = E_x\left[\frac{i}{c}\frac{k_y^2}{k_0} - \varepsilon_1\frac{ik_0}{c}\right] + E_y\left[\frac{i}{c}\frac{k_x k_y}{k_0} + \varepsilon_2\frac{k_0}{c}\right] \tag{7}$$

$$\frac{dE_x}{dz} = -\frac{ik_x k_y c}{k_0\varepsilon_3}b_x + b_y\left[\frac{ik_x^2 c}{k_0\varepsilon_3} - ik_0 c\right] \tag{8}$$

$$\frac{dE_y}{dz} = b_x\left[\frac{-ik_y^2 c}{k_0\varepsilon_3} + ik_0 c\right] + b_y\left[\frac{ik_x k_y c}{k_0\varepsilon_3}\right] \tag{9}$$

$$b_z = \frac{1}{k_0 c}\left[k_y E_x - k_x E_y\right] \tag{10}$$

$$E_z = \frac{c}{k_0\varepsilon_3}\left[k_x b_y - k_y b_x\right] \tag{11}$$

where ε_1, ε_2 and ε_3 are associated with the Pedersen, Hall and direct conductivities respectively and $k_0 = \omega(\mu_0\varepsilon_0)^{1/2}$.

27.2.1 Boundary conditions

To obtain a solution to the differential Equations (6)–(9) we require four boundary conditions. The bottom boundary conditions allow for finite ground conductivity where the ULF wave decays in a uniform, isotropic medium. The bottom boundary conditions are therefore,

$$\frac{\delta E_x}{\delta x} - \alpha E_x = 0 \tag{12}$$

$$\frac{\delta E_y}{\delta y} - \alpha E_y = 0 \tag{13}$$

where α represents the skin depth.

The upper boundary assumes an interface between the ionospheric plasma and uniform, ideal MHD conditions. It is then possible to specify various mixtures of fast and shear Alfvén mode incident energy. For the shear Alfvén mode (in ideal MHD) there is no magnetic perturbation in the direction of the background magnetic field. Therefore, from Equation (2)

$$b_z = \frac{i}{\omega}\left(k_x E_y - k_y E_x\right) = 0 \quad . \tag{14}$$

A pure fast mode wave must satisfy $\nabla\cdot\mathbf{E} = 0$ (Cross, 1988). For ideal MHD, the field aligned electric field is zero ($E_Z = 0$) so that

$$k_x E_x + k_y E_y = 0 \quad . \tag{15}$$

If the energy is a mixture of these two distinct MHD modes then we can resolve the perturbation electric field of the wave into fast and shear Alfvén mode components as

$$E_{mix} \propto \alpha\left(k_x, k_y, 0\right) + \beta\left(-k_y, k_x, 0\right) \tag{16}$$

where α and β are complex constants that represent the amount each distinct mode contributes to the total incident ULF energy.

To set the upper boundary condition for a given mixture of fast and shear Alfvén mode waves and a specified horizontal structure (k_x and k_y), we choose α and β and one of the perturbed electric field components, say E_y. The other component, E_x, is then fixed by

$$\frac{E_x}{E_y} = \frac{\alpha k_x - \beta k_y}{\alpha k_y + \beta k_x}. \tag{17}$$

The system can now be solved for the ULF wave fields provided the medium is specified through the conductivity tensor as described in Zang and Cole (1994). The atmospheric and ionospheric parameters were obtained from the MSIS-86 and IRI-90 models for the day 12 January 1994.

27.3 HF waves in the ionosphere

Considering the properties of the ionosphere and the frequency range of interest (HF), Appleton (1927) developed an equation that describes how the refractive index, n, varies with frequency, ω_{HF}. For vertical incidence,

$$n^2 = 1 - \frac{X}{(1 - iZ) + iY_L R} \tag{18}$$

where

$$R = \frac{i}{2Y_L}\left(\frac{Y_T^2}{1 - X - iZ} \pm \sqrt{\frac{Y_T^4}{(1 - X - iZ)^2} + 4Y_L^2}\right) \tag{19}$$

$$X = \frac{N_e e^2}{\varepsilon_0 m_e \omega_{HF}^2} \tag{20}$$

$$Y_{L,T} = \frac{eB_{L,T}}{m_e \omega_{HF}} \tag{21}$$

and m_e is the electron mass, e is the electron charge (negative), $Z = \nu/\omega_{HF}$ for ν the electron collision frequency, N_e is the electron number density, and $B_{L,T}$ is the longitudinal and transverse components of the magnetic field relative to the HF wave propagation vector. The electron collision term gives a complex refractive index where $n = \mu + i\chi$ (Budden, 1985). For frequencies used for sounding the ionosphere, Z is usually very small and the refractive index is real so that (18) becomes

$$\mu^2 = 1 - \frac{2X(1 - X)}{2(1 - X) - Y_T^2 \pm \sqrt{Y_T^4 + 4(1 - X)^2 Y_L^2}} \tag{22}$$

The ± sign identifies the ordinary and extraordinary propagation modes.

27.4 Doppler shifts in HF signals due to ULF waves

The SP model identifies three mechanisms that contribute to Doppler shifts in HF signals due to ULF wave energy. For a coordinate system

where x is positive northward, y is positive eastward and z is positive down, the Doppler velocity from the magnetic mechanism is

$$V_1 = -i\omega \int_0^{Z_R} \left[\frac{\delta\mu}{\delta B_L} \frac{\delta B_L}{\delta t} + \frac{\delta\mu}{\delta B_T} \frac{\delta B_T}{\delta t} \right] dz \qquad (23)$$

where the magnetic field is $\mathbf{B} = \mathbf{B}_L + \mathbf{B}_T$, the longitudinal and transverse components respectively and the ULF magnetic field varies as $\mathbf{B} = \mathbf{B}_0 + \mathbf{b}_0 e^{-i\omega t}$. Equation (23) describes the change in refractive index due to magnetic field variations from the ULF wave and shows that the refractive index can change from mechanisms other than electron motion.

The advection mechanism involves the electron density, N and is given as

$$V_2 = -\int_0^{Z_R} \left[\frac{\delta\mu}{\delta N} (v \cdot \nabla N) \right] dz \quad . \qquad (24)$$

This describes the vertical bulk motion of electrons driven by the electric field of the ULF wave and is essentially the same as the first mechanism described by Rishbeth and Garriott (1964).

The compression mechanism is

$$V_3 = -\int_0^{Z_R} \left[\frac{\delta\mu}{\delta N} N(\nabla \cdot v) \right] dz \qquad (25)$$

which describes the Doppler shift due to the compression and rarefaction of the plasma due to the ULF wave energy. A fourth mechanism identified by Poole *et al.* (1988) is zero since electron production and loss terms are assumed to be equal. The Doppler velocity was calculated from Equations (23)–(25) as $V^* = V_1 + V_2 + V_3$.

27.5 Experimental data

The Doppler sounder technique monitors the frequency shift of a HF carrier reflected from ionospheric heights. During the southern hemisphere summer of 1994, the VNG signal from Llandillo (near Sydney) was received at Clarencetown, 30 km north of Newcastle using modified commercial Kenwood R-5000 transceivers capable of detecting frequency variations of < 0.1 Hz (Menk *et al.*, 1995; Marshall, 1996). The ULF data were obtained using induction magnetometers (Fraser *et al.*, 1991) comprising two perpendicular sensors oriented geomagnetic north-south and east-west. Data were recorded every two seconds using a PC controlled 12 bit A/D card and stored on disk. Time accuracy was typically < 10 ms. Extensive descriptions of the data and analysis details are given in Marshall (1996) and Dunlop (1998).

In this paper, we concentrate on modelling the Doppler shifts in HF due to ULF wave energy around a FLR frequency. Typical results for the amplitude-ratio and cross-phase between data obtained from the magnetometer north-south sensor and the Doppler receiver are shown in Figure 1. The FLR frequencies were identified by an independent technique using two magnetometers (e.g. Waters *et al.*, 1991) and

correspond with the maxima in the amplitude-ratio and dips in cross-phase. The SP models only model the shear Alfvén ULF mode which is dominantat the FLR frequencies. Figure 1(b) shows this mode has a narrow bandwidth at the two harmonics identified by the arrows, while most of the spectrum exhibits a different and almost constant cross-phase value. The modelling will concentrate on reproducing the observed amplitude and phase across the FLR at ~ 50 mHz.

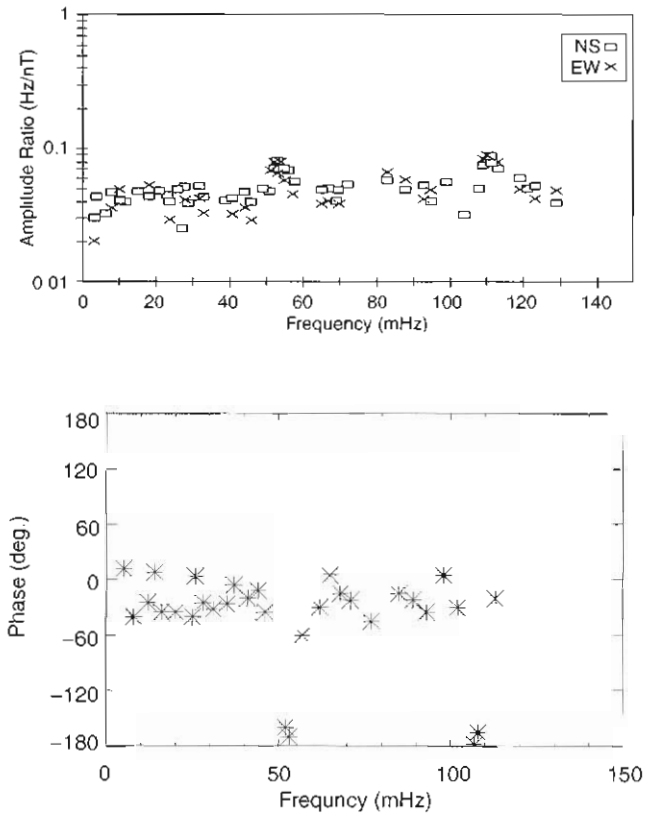


Figure 1. *Comparison of ionospheric Doppler shift and north-south magnetometer data. (a) Amplitude ratio and (b) cross phase for data recorded at L = 1.8 on 12 January 1994.*

27.6 Data from the model

The first step for the modelling is to obtain the variation of the ULF electric and magnetic perturbation fields with height through the ionosphere/atmosphere system. The excitation of FLR's arises from inward propagating fast mode energy from the magnetosphere. At the location (magnetic field shell) where the frequencies match, the FLR grows and narrows in spatial extent (e.g. Tamao, 1966). Choosing a frequency range of 40 to 60 mHz, both the ULF wave mode mixture

and spatial structure change as the frequency passes through the FLR at 50 mHz. A number of measurements have been made of ULF azimuthal wavenumbers (e.g. Ansari and Fraser, 1986). Based on these, we set a pure fast mode at both 40 and 60 mHz with $k_x = k_y = 8 \times 10^{-8}$ m^{-1}. At 50 mHz we set 90% shear Alfvén mode and increase k_x to 1×10^{-7} m^{-1}. The values for mode mix and k_x at the other frequencies were altered using a suitable Gaussian fit joining these three points.

The amplitude and phases for the ULF perturbation magnetic fields for a fast mode at 40 mHz and an Alfvén mode at 50 mHz are shown in Figure 2. It is important to keep in mind that the ULF energy obeys ideal MHD only at the topside boundary. Through the ionosphere, the Maxwell equations are solved and the medium exhibits finite, anisotropic conductivity. For the given horizontal spatial structure, the 40 mHz fast mode horizontal magnetic perturbations, b_x and b_y, pass through the ionosphere without much change in amplitude or phase. The shear Alfvén case at 50 mHz shows a decrease in amplitude and

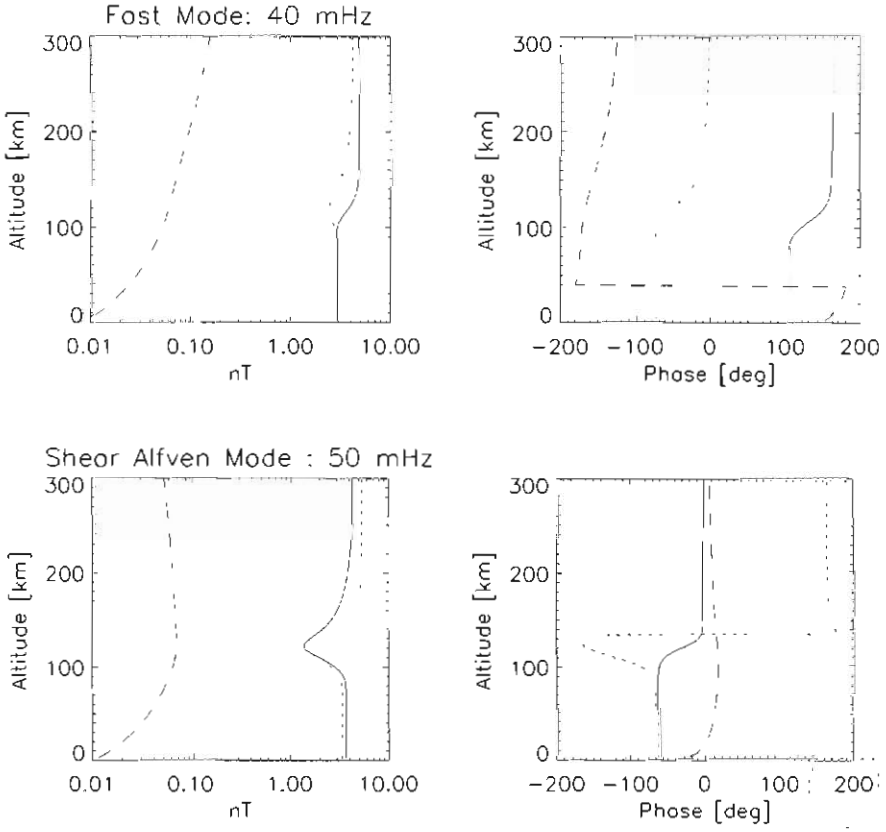


Figure 2. *The ULF magnetic field perturbations through the modelled ionosphere. The absolute value and phase are shown where b_x is short dashed, b_y is long dashed and b_z is dash-dot lines. (a) Fast mode wave at 40 mHz and $k_x = k_y = 8 \times 10^{-7}$ m^{-1} (b) 90% shear Alfvén mode for $k_x = 1 \times 10^{-6}$ m^{-1} and $k_y = 8 \times 10^{-7}$ m^{-1}.*

large shift in phase in both b_x and b_y around 120 km altitude. This is consistent with 'ionospheric shielding' described by Hughes (1974). This may be explained using the Maxwell equation where no currents flow in the neutral atmosphere so that

$$J_z = \frac{1}{\mu_0}(\nabla \times b) = \frac{1}{\mu_0} k_\perp \times b_{horiz} = 0 \quad . \tag{26}$$

This means that the horizontal components of the perturbation magnetic field must be either zero or parallel to $k_\perp$ which gives rise to a rotation of the $b_\perp$ components by 90°. This effect is usually shown with b_y decreasing to zero at ground level (e.g. Hughes and Southwood, 1976). However, this only occurs when the shear Alfvén wave at the top boundary has the magnetic perturbation mostly in the y component, i.e. for $k_x >> k_y$. The wave polarisation is determined by the MHD conditions for the shear Alfvén mode ($k \cdot E \neq 0$) and the horizontal spatial structure. The shielding still occurs for different wave polarisations while the magnitude of b_y on the ground is determined by the projection of the polarisation ellipse on the coordinate axes.

The Doppler shifts and phases as a function of height at 2 km intervals for the shear Alfvén mode case are shown in Figure 3. The major contribution to the Doppler shift comes from mechanism V_3 which describes changes in the refractive index due to the divergence of the electron velocity at these altitudes. The velocities have been calculated from the perturbation electric fields of the ULF energy and the electron mobilities as described by Sutcliffe and Poole (1989). The increase in Doppler shift from around 200 km and higher, and the relatively constant phase with height is consistent with the results of Sutcliffe and Poole (1989). A direct comparison is not possible as; (i) the ionospheric parameters are different, (ii) the ULF wave fields are for incident shear Alfvén mode waves but the spatial structure parameters are different and

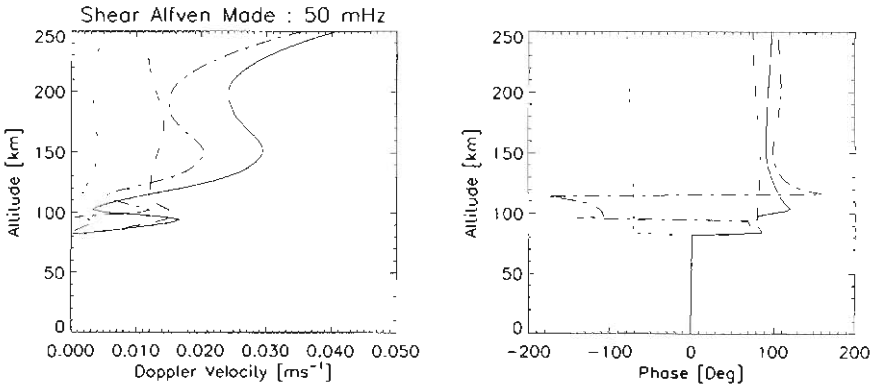


Figure 3. *Absolute values and phases of the Doppler shifts due to the three mechanisms (see text) for ULF energy with 90% shear Alfvén mode.*

(iii) Sutcliffe and Poole (1989) used a main magnetic field inclined at 60° while ours is almost vertical.

The Doppler shifts associated with incident fast mode ULF energy at 40 mHz are shown in Figure 4. Mechanism V3 is still the largest contributor while the relative contribution of mechanism V2 (bulk vertical motion of electrons) has decreased. The fast and shear Alfvén modes show similar contributions to the Doppler velocity from mechanism V1. This describes Doppler velocities due to changes in the magnetic field strength. Like the sheal Alfvén case, the phases are relatively constant in the F region. However, for HF reflected in the E region, the phase becomes quite sensitive to the medium properties. Finally, the relative contributions from V1 and V2 in Figure 4 are altitude dependent with V1 larger at higher compared with V2 larger at lower altitudes.

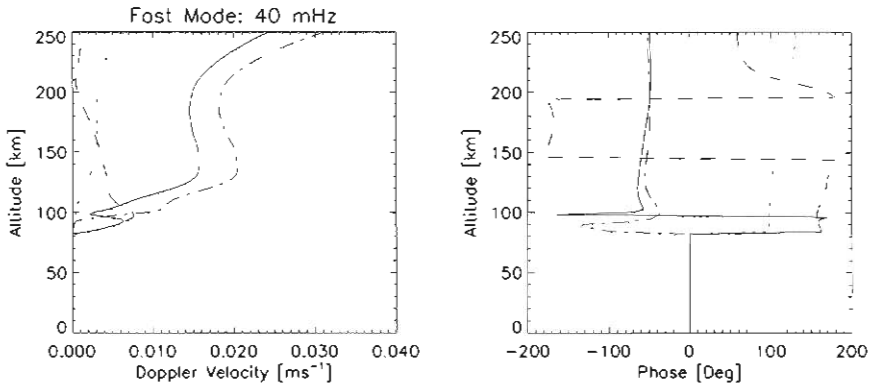


Figure 4. *Absolute values and phases of the Doppler shifts due to the three mechanisms (see text) for the 40 mHz fast mode ULF energy.*

27.7 Discussion

Figure 1 shows experimental values for amplitude-ratios and phase differences calculated from data recorded by the Doppler sounder and a magnetometer located beneath the reflection region. The Doppler receiver was tuned to 5 MHz. Assuming near vertical incidence and the ionospheric parameters for 12 January 1994, this corresponded with a reflection height of around 200 km. The larger Doppler shift for the shear Alfvén mode (FLR) compared with the fast mode incident energy is consistent with the ratio results in Figure 1. In order to compare the ratio and phase difference results with the model, various ULF mode mixtures at the top boundary were used and the associated Doppler shifts were calculated as described above. The data at 200 km were obtained

from the model and the results are shown in Figure 5. The dip in phase difference is reproduced although not to the extent shown in Figure 1. The phase differences are sensitive to the choice of k_x and k_y. There are quite a few measurements of k_y at low-latitudes and we have used a typical value. However, the choice for k_x is more guesswork. The difficulty lies in measuring the phase difference at latitudinally spaced magnetometer sites at frequencies which are resonant. We are investigating ways of obtaining accurate measurements of k_x.

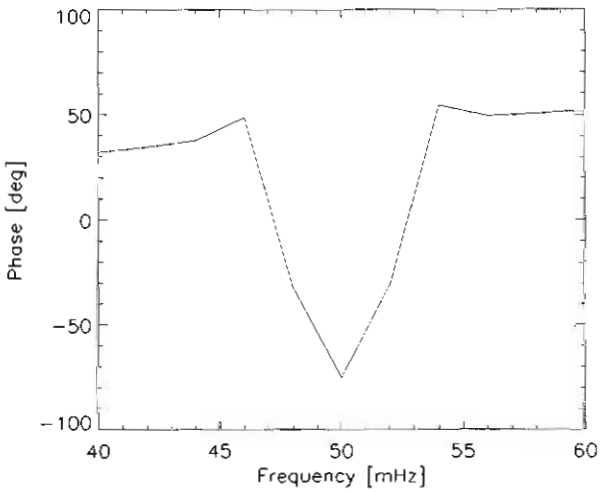


Figure 5. *Cross phase between the total Doppler velocity at 200 km altitude and the value of bx at the ground calculated from the model.*

Doppler oscillations in HF signals may arise from other processes such as gravity waves, acoustic modes and other phenomena that alter the electron density. A further complication arises from the birefringent medium which gives the ordinary and extraordinary HF modes. These two modes at a given frequency reflect at different altitudes and may interfere (beat) at the receiver. All these effects make the role of the magnetometer vital to identifying Doppler shifts in HF due to ULF wave energy. The study by Marshall (1996) showed that for intervals of data that showed clear correspondence between the Doppler sounder and magnetometer data, the ordinary mode was around 20 dB larger compared with the extraordinary mode. Therefore the modelling in this paper has only considered the ordinary HF mode.

The modelling for the ULF signal through the ionosphere assumed a vertical magnetic field. This allowed the definition of various ULF wave mode mixtures at the top side boundary. The difficulty with the vertical

field configuration comes from Equation (21) for vertical incidence of the HF signal where Y_T goes to zero. What actually happens is that we encounter a radio window (Budden, 1985). The condition for wave reflection for HF propagation along the magnetic field only involves the extraordinary mode. If a wave propagates in the ordinary mode along the field through a medium where the refractive index decreases with distance, then at the location where reflection would occur (the window point) if the field and HF propagation vector were not parallel an extraordinary wave appears. To avoid this problem, a small B_T (2% of B_L) was included in the calculations for the Doppler shifts. It would also be desirable to give the main field a realistic inclination to model the data recorded at Australian latitudes. We are presently working on formulating the boundary conditions required to retain the flexibility of being able to specify the ULF wave mode mixture for a non-vertical field.

27.8 Conclusion

HF signals that propagate via the ionosphere are known to exhibit Doppler oscillations caused by ULF wave energy incident from the magnetosphere. In this paper we have extended previous models of ULF wave propagation through the ionosphere so that the ULF wave mode mixture may be specified. This allows a greater range of the ULF spectrum and the effects on HF Doppler shifts to be investigated. We have found that the vertical bulk motion of the electrons is a minor contributor to the Doppler shifts with the major contribution arising from the divergence of the electron velocities. For fast mode energy, the role of the compression mechanism becomes larger. Future developments of the model include a formulation for non-vertical magnetic field, improving the estimate for k_x and comparing with Doppler sounder and HF radar data.

References

Ansari, I.A. and Fraser, B.J. (1986). A multistation study of low latitude Pc3 geomagnetic pulsations. *Planetary and Space Science*, 34: 519–536.

Budden, K.G. (1985). *The Propagation of Radio Waves*. Cambridge University Press, Cambridge.

Chen, L. and Hasegawa, A. (1974). A theory of long period magnetic pulsations 1. Steady state excitation of field line resonances. *Journal of Geophysical Research*, 79: 1024–1032.

Cross, R. (1988). *An Introduction to Alfvén Waves*. IOP Publishing, Bristol, England.

Dunlop, I.S. (1998). *Observations and modelling of hydromagnetic waves in the ionosphere*. PhD. Thesis, The University of Newcastle, NSW, Australia.

Fraser, B.J., McNabb, P.W., Menk, F.W. and Waters, C.L. (1991). A personal computer induction magnetometer system for recording geomagnetic pulsations. In: Burns, G. (Ed.). *ANARE Research Notes 80*. Australian Antarctic Division. Pp. 83–92.

Harang, L. (1939). Pulsations in an ionized region at height of 650–800 km during the appearance of giant pulsations in the geomagnetic records. *Journal of Terrestrial Magnetism and Atmospheric Electricity*, 44. p. 17.

Hughes, W.J. (1974). The effect of the atmosphere and ionosphere on long period magnetospheric micropulsations. *Planetary and Space Science*, 22: 1157–1172.

Hughes, W.J. and Southwood, D.J. (1976). An illustration of modification of geomagnetic pulsation structure by the ionosphere. *Journal of Geophysical Research*, 81: 3241–3247.

Jacobs, J.A. and Watanabe, T. (1966). Doppler frequency changes in radio waves propagating through a moving ionosphere. *Radio Science*, 1: 257–264.

Marshall, R.A. (1996). *Observations of hydromagnetic waves in the ionosphere*. PhD. Thesis. The University of Newcastle, NSW, Australia.

Menk, F.W., Marshall, R.A., McNabb, P.W. and Dunlop, I.S. (1995). An experiment to study the effects of geomagnetic fluctuations on ionospheric HF radio paths. *Journal of Electricity and Electrical Engineering*.

Pitteway, M.L.V. (1965). The numerical calculation of wave fields, reflection coefficients and polarizations for long radio waves in the lower ionosphere. Royal Society Philosophical Transactions of the Royal Society, 257. p. 243.

Poole, A.W.V. and Sutcliffe, P.R. (1987). Mechanisms for observed total electron content pulsations at mid latitudes. *Journal of Atmospheric and Terrestrial Physics*, 49: 231–236.

Poole, A.W.V., Sutcliffe, P.R. and Walker, A.D.M. (1988). The relationship between ULF geomagnetic pulsations and ionospheric Doppler oscillations: derivation of a model. *Journal of Geophysical Research*, 93: 14656–14664.

Rishbeth, H. and Garriott, O.K. (1964). The relationship between simultaneous geomagnetic and ionospheric oscillations. *Radio Science*, NBS, 68D: 339–343.

Southwood, D.J. (1974). Some features of field line resonances in the magnetosphere. *Planetary and Space Science*, 22: 483–491.

Sutcliffe, P.R. and Poole, A.W.V. (1989). Ionospheric Doppler and electron velocities in the presence of ULF waves. *Journal of Geophysical Research*, 94: 13505–13514.

Sutcliffe, P.R. and Poole, A.W.V. (1990). The relationship between ULF geomagnetic pulsations and ionospheric Doppler oscillations: model predictions. *Planetary Space Science*, 38: 1581–1589.

Tamao, T. (1966). Transmission and coupling resonance of hydromagnetic disturbances in the non-uniform Earth's magnetosphere. *Science Report*. Tohoku University, Series 5, 17. p. 43.

Watermann, J. (1987). Observations of correlated ULF fluctuations in the geomagnetic field and in the phase path of ionospheric HF soundings. *Journal of Geophysical Research*, 61: 39–45.

Waters, C.L., Menk, F.W. and Fraser, B.J. (1991). The resonance structure of low latitude Pc3 geomagnetic pulsations. *Geophysical Research Letters*, 18: 2293–2296.

Yumoto, K., Saito, T., Akasofu, S-I., Tsurutani, B.T. and Smith, E.J. (1985). Propagation mechanism of daytime Pc3-4 pulsations observed at synchronous orbit and ground-based stations. *Journal of Geophysical Research*, 90: 6439–6450.

Zhang, D.Y. and Cole, K.D. (1994). Some aspects of ULF electromagnetic wave relations in a stratified ionosphere by the method of boundary value problem. *Journal of Atmospheric and Terrestrial Physics*, 56: 681–690.

Zhang, D.Y. and Cole, K.D. (1995). Formulation and computation of hydromagnetic wave penetration into the equatorial ionosphere and atmosphere. *Journal of Atmospheric and Terrestrial Physics*, 57: 813–819.

28. COORDINATED OBSERVATIONS OF FORCED AND RESONANT FIELD LINE OSCILLATIONS AT HIGH-LATITUDES

F.W. Menk[1], T.K. Yeoman[2], D. Wright[2] and M.E. Lester[2]

[1] Department of Physics
Cooperative Research Center for Satellite Systems
University of Newcastle
Callaghan NSW 2308 Australia
(email: physpuls3@ cc.newcastle.edu.au)

[2] University of Leicester
Leicester LE1 7RH
United Kingdom

Abstract

In this paper we examine the geomagnetic signatures of solar wind impulses using magnetometer array data, a bistatic HF radar, and a co-located imaging riometer. The impulses were associated with 1.6 mHz magnetic pulsations that were observed across a wide range of latitudes. These were most likely due to field line oscillations driven by incoming fast mode waves. At the same time, the radars recorded oscillatory velocity features in ground backscatter returns consistent with ionospheric motions driven by the downgoing waves. Accompanying decreases in cosmic noise absorption were evident in the imaging riometer data.

28.1 Introduction

The purpose of this paper is to examine the magnetic and ionospheric signatures of solar wind pressure perturbations. Geomagnetic activity is driven by energy and mass transfer from the solar wind, and the outermost sunward field lines map most directly to the wave and particle entry regions. The injected particles may give rise to features of dayside auroras (Lui and Sibeck, 1991), while the waves may propagate large distances through the magnetosphere and couple to field line resonances and global cavity modes (Warnecke *et al.*, 1990).

High-latitude data sets are often characterised by the appearance of transient features, but the interpretation of these is controversial. Discussion in the literature has focused on radial displacements of the magnetopause due to perturbations in the solar wind dynamic pressure (Farrugia *et al.*, 1989; Sibeck, 1990) and impulsive plasma transfer and magnetic field reconnection events across the magnetopause (Lemaire, 1977; Russell and Elphic, 1979; Saunders *et al.*, 1984). Kelvin-Helmholtz surface waves driven by the velocity shear at

the magnetopause may also produce similar signatures in ground magnetometer records (Pu and Kivelson, 1983). There is a need for detailed case studies in order to distinguish the causative mechanism. An important aspect of this is the intercalibration of signatures from different instruments.

In this paper we report the observation of high-latitude transient events recorded on a selected day on the ground with the IMAGE and SAMNET magnetometer arrays, and simultaneously in the ionosphere with the CUTLASS HF radar and an imaging riometer. The events were stimulated by solar wind pressure perturbations and we thus identify the magnetospheric response to this phenomenon.

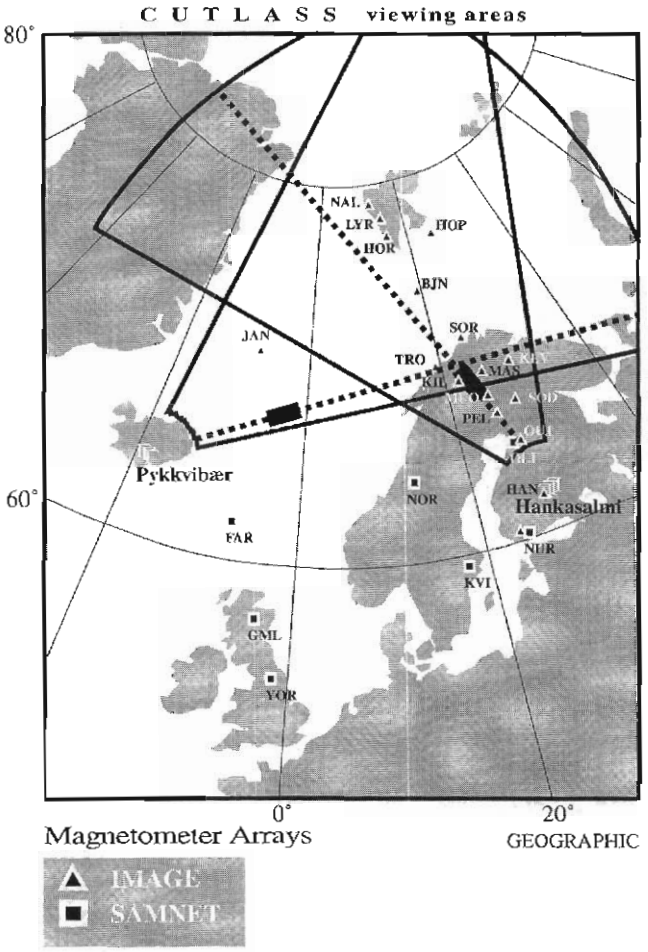


Figure 1. *Map in geographic coordinates showing IMAGE and SAMNET magnetometer stations and the CUTLASS beams of interest: Finland beam 6 and Iceland East beam 14 (broken lines). The IRIS imaging riometer was located at KIL. Shaded rectangular blocks denote regions in the ionosphere from which Doppler oscillations were recorded.*

28.2 Data sources

28.2.1 Ground magnetometers

Magnetic field observations were obtained from the IMAGE and SAMNET arrays spanning Scandinavia and Svalbard. Station locations are shown in Figure 1, where the triangles represent the IMAGE stations used here and the squares indicate the SAMNET locations. The IMAGE (International Monitor for Auroral Geomagnetic Effects) array comprises a coordinated multinational network of fluxgate magnetometers sampling the geographic X, Y and Z components of the geomagnetic field each 10 s with a resolution of 0.1–1 nT (Lühr, 1994; Lühr *et al.*, 1998). For use here data were rotated into the geomagnetic H, D and Z reference frame; see Howard *et al.* (1999) for further details. Corrected geomagnetic coordinates (CGM; http://nssdc.gsfc.nasa.gov/space/cgm/cgm.html) are used throughout this paper.

SAMNET (UK Sub-Auroral Magnetometer NETwork) is operated by the University of York, U.K., and comprises fluxgate magnetometers that during the interval considered here sampled the geomagnetic H, D and Z components each 5 s with a resolution of 0.25 nT. Further details on SAMNET appear in Yeoman *et al.* (1990).

The IMAGE and SAMNET data used in this paper were obtained from the respective home pages, from which further details, including station coordinates, are available. These addresses are:

http://www.geo.fmi.fi/image/ and
http://samsun.york.ac.uk/samnet_home.html.

Both magnetometer arrays are being continually upgraded, adding more stations and improving timing accuracy. For example, SAMNET now samples each 1 s and uses GPS timing.

28.2.2 HF radar

The ionosphere above northern Scandinavia and Svalbard can be studied with a number of instruments. One of these is the CUTLASS (Cooperative UK Twin Located Auroral Sounding System) bistatic HF coherent radar, with stations at Pykkvibær in Iceland and Hankasalmi in Finland (see Figure 1). The former is called the CUTLASS Iceland East radar, and since it looks across the magnetic meridian is used to identify features moving azimuthally. The Finland radar looks toward the magnetic pole and accordingly is better suited for examining equatorward or poleward moving features. Further details on CUTLASS are available from http://ion.le.ac.uk/cutlass/cutlass.html.

CUTLASS is a component of the international SuperDARN HF radar network (Greenwald *et al.*, 1995). In standard operating mode these radars sweep over a 52° azimuth sector using 16 evenly spaced beams. Each beam is gated into up to 75 range bins with spatial resolution of typically 45 km, giving a total field-of-view of order 3×10^6 km^2. In standard mode, the integration time for each beam position is 7 s, and the cycle time 2 minutes.

The operation of HF radars was described by Greenwald *et al.* (1985). The radars detect echoes that are backscattered from decameter-scale field-aligned ionospheric plasma density irregularities. These are produced in the high-latitude F-region by plasma drifts and density gradients (e.g. Fejer and Kelley, 1980; Tsunoda, 1988) and drift with the ambient plasma motion (Villain *et al.*, 1985). In each beam the HF radars measure the line-of-sight Doppler velocity, the backscattered power and the spectral width of the echo from these structures.

In addition to operation in standard modes, each SuperDARN radar also operates in discretionary modes for specific studies. The observations reported here were obtained when CUTLASS was operating in a non-standard scan mode optimised for high resolution studies of a specific target region. In this mode three adjacent beams are scanned with the central beam scanned again between each set. The Finland radar thus scanned beams 6, 7, 6, 5, 6, 7, 6, ... , and similarly the Iceland East radar scans were centered on beam 14. These beams were sampled each 7 s and the first range gate was set at a distance of 180 km. The radio frequency was in the range 9.900–10.000 MHz. The central beams are represented in Figure 1 by broken lines, and their intersection region was in the F-region just east of Tromsø. The solid lines in Figure 1 represent the standard fan-shaped SuperDARN coverage.

An important feature of radars such as CUTLASS is measurement of the elevation angle of the backscatter returns. This is achieved using an interferometric technique and can provide considerable assistance in discriminating the direction of ground scatter and the altitude of irregularity structures (Milan *et al.*, 1997, 1997b).

28.2.3 Imaging riometer

Information on the magnitude and structure of D-region absorption for the same location probed by CUTLASS was obtained from the IRIS (Imaging Riometer for Ionospheric Studies) experiment operated at Kilpisjärvi (KIL), Finland, by the University of Lancaster, U.K. IRIS examines ionospheric absorption of incoming galactic radio noise at a frequency of 38.2 MHz with 49 antenna beams in a 7 x 7 array and a sampling time of 1 s. The total field-of-view at 90 km altitude is of order 240 x 240 km^2. The experiment was described in detail by Browne *et al.* (1995), and data were obtained from: http://www.dcs.lancs.ac.uk/iono/iris/. The location of KIL is also shown in Figure 1.

28.2.4 Satellite data

Solar wind data are obtainable from the WIND spacecraft. For the time of interest here this was located in the upstream solar wind near GSE (x, y, z) = (172.5, -5.6, 11.23) R_E. We used solar wind and magnetic field data from the SWE (Solar Wind Experiment) and MFI (Magnetic Fields Investigation) instruments respectively. Further information on WIND and its experiments is available from

http://www-spof.gsfc.nasa.gov/istp/wind/, while the data used here may be accessed at http://cdaweb.gsfc.nasa.gov/.

General information on magnetospheric topology was obtained by reference to ion and electron flux data from the DMSP F12 spacecraft. This is one of a series of spacecraft in an approximately 830 km altitude Sun-synchronous, 101 minute period polar orbit. Spectra of low energy ion and electron fluxes at high-latitudes are produced by the SSJ/4 electrostatic analyser instruments. Full details of the orbit and data, including spectrograms, are available at http://sd-www.jhuapl.edu/aurora/; see also http://web.ngdc.noaa.gov/dmsp/dmsp.html. The use of DMSP particle spectra to determine the magnetospheric boundaries was discussed by Newell and Meng (1988, 1992) and Newell *et al.* (1989, 1991).

28.3 Results

28.3.1 Magnetic pulsations and solar wind observations

In this paper we focus on observations from 23 February 1996, when Pc5 type magnetic pulsations were recorded across the IMAGE and SAMNET magnetometer arrays, and corresponding Doppler oscillations occurred in the CUTLASS records. These pulsations lasted for several hours, during which K_p was 3^+.

Figure 2 shows stacked H component time series plots of the magnetometer data from selected IMAGE and SAMNET stations for the interval 0900–1300 UT on this day, bandpass filtered between 1 mHz and 40 mHz. The D component plots are similar and therefore not shown. Higher resolution plots have also been inspected but for brevity are not presented here. Several features are apparent in the plots.

(i) A large, transient bipolar event commenced at 0932 UT at all stations.

(ii) This was followed for some hours by wave-like activity that was remarkably similar at all stations < 74° latitude, although amplitudes increased with increasing latitude.

(iii) Signals were somewhat different at latitudes ≥ 74°.

(iv) Shorter period pulsations were superimposed on the large, long period variations at all stations.

We now consider each of these features in turn.

WIND observations are presented in Figure 3, and show that the solar wind speed was in the range V = 400–430 km s^{-1} throughout this time. This includes a small, sharp increase from ~ 410 km s^{-1} to ~ 430 km s^{-1} at 0850 UT. More importantly, the ion density decreased from n ~ 20 cm^{-3} at 0841 UT to 13 cm^{-3} at 0852 UT and 6 cm^{-3} at 0855 UT. This represents a decrease in solar wind ram pressure, nV^2, by a factor of 2–3. Given the upstream location of WIND, this negative pressure pulse should reach the magnetopause around 0930–0935 UT. The corresponding WIND IMF

data show B_z turning negative around 0838 UT, with B_y increasing from -6 to +5 nT over 0845–0852 UT.

There are further pressure perturbations at WIND after 0950 UT and in particular near 1130 UT (decrease by ~ 25%), 1140 UT (increase by ~ 50%), 1150 UT (decrease by ~ 7%), 1200 UT (decrease by ~ 17%). These perturbations exhibit 10 minute periodicity (~ 1.7 mHz frequency). A large positive excursion in B_z and B_x between about 1140 UT and 1150 UT at WIND may be connected with the sudden decrease in pulsation activity at the ground between 1230 UT and 1244 UT.

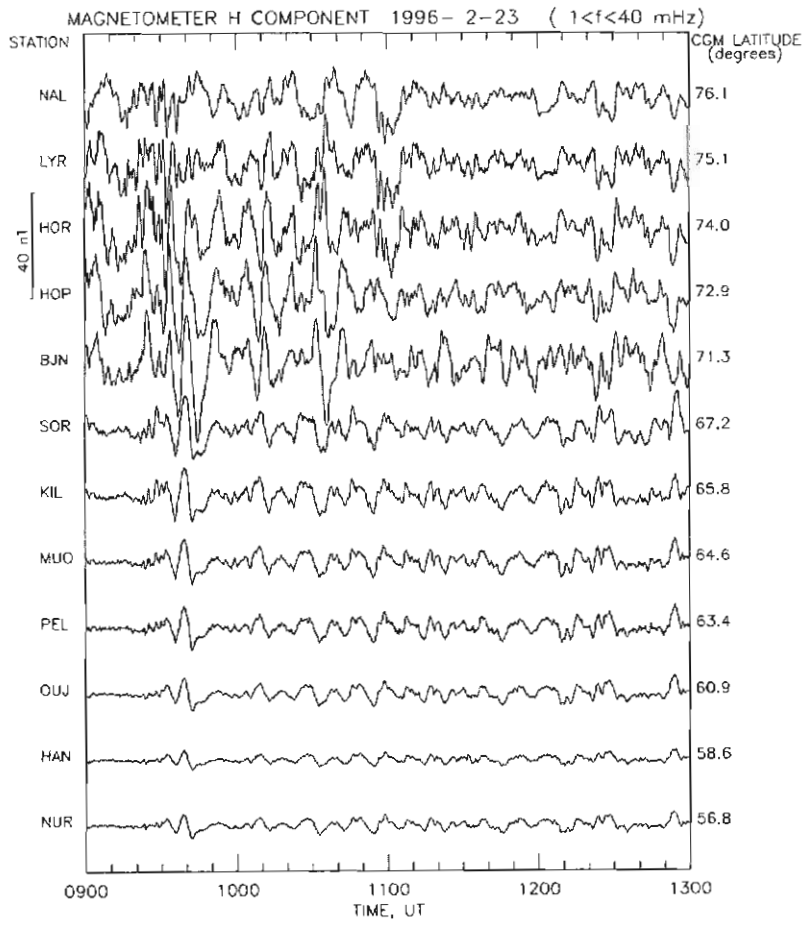


Figure 2. *Stacked H component time series for selected IMAGE and SAMNET stations, 0900–1300 UT, 23 February 1996.*

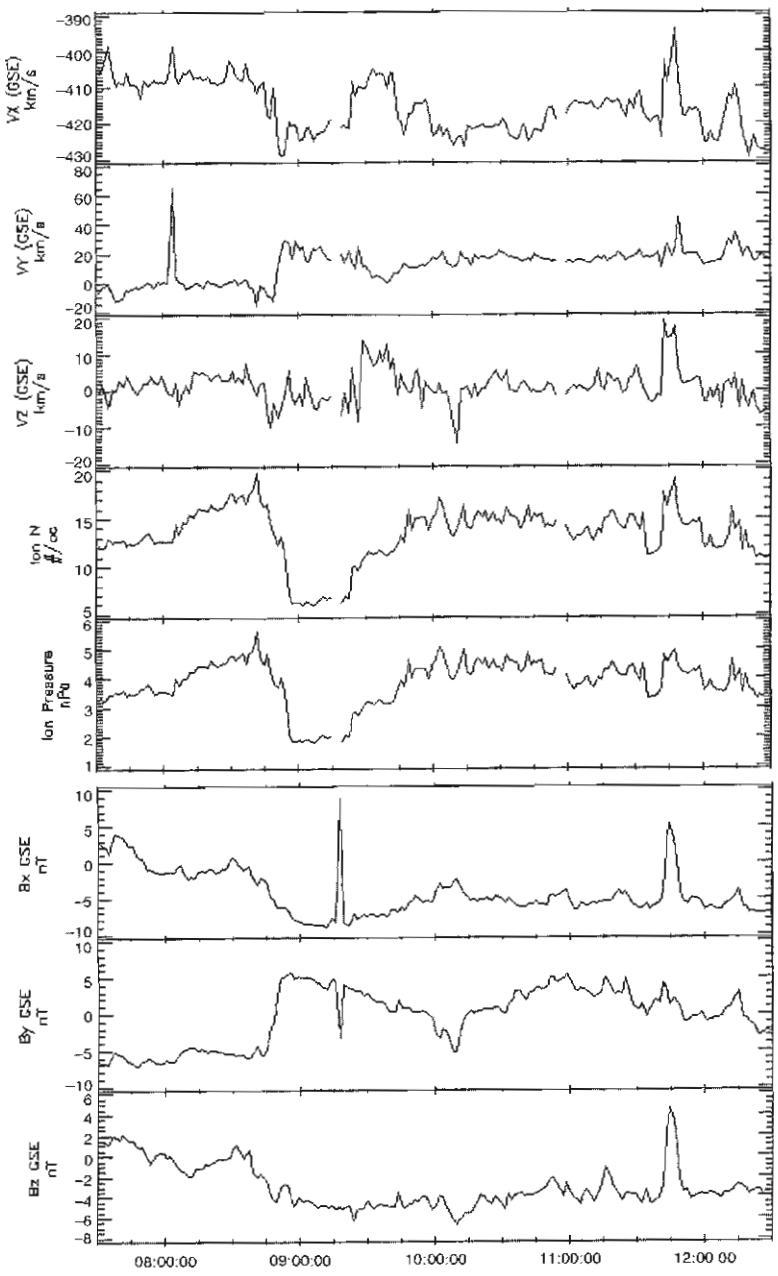


Figure 3. *Upstream WIND spacecraft data, 0730–1030 UT, 23 February 1996. Upper three panels show solar wind velocity, then ion density, ion pressure, and the IMF.*

Pressure pulses of the magnitude observed here are believed to be fairly common (Sibeck, 1990) and have been associated with ringing type magnetic pulsations on the ground with periods of a few minutes (Takahashi *et al.*, 1988; Farrugia *et al.*, 1989; Sibeck, 1990), including field line resonances (Potemra *et al.*, 1989; Warnecke *et al.*, 1990; Parkhomov *et al.*, 1998; Prikryl *et al.*, 1998). Accordingly, it seems likely the bipolar event seen in the magnetometer records at 0932 UT is due to the negative solar wind pressure impulse. This is followed by some hours of similar pulsation activity across a wide range of latitudes.

Stacked power spectra for the ground magnetometer H components are presented for a representative interval in Figure 4. The spectra have been normalised and weighted by $f^{1.0}$ to better illustrate higher frequency features. Pulsations at a frequency of 1.6 mHz dominate at all but the two highest latitude stations. A secondary peak around 8 mHz also occurs at many stations.

We now consider the activity at the highest latitude stations, $\geq 74°$ CGM latitude (LYR and NAL). The DMSP F13 spacecraft passed over the arctic region north of Scandinavia around 0847 UT on this day. Particle energies and fluxes characteristic of the cusp were seen around 80° geomagnetic latitude, and boundary layer particles just equatorward of this. Therefore, the magnetometer signals at the highest latitude IMAGE stations are probably associated with closed auroral oval field lines. This is confirmed by the radar observations presented in the next section. It is therefore not surprising that the magnetometer signals at these latitudes appear more noisy and less ordered than at lower latitudes.

The field line eigenfrequency was determined from the cross-phase between pairs of adjacent IMAGE stations, using the method described by Waters *et al.* (1995). The variation in eigenfrequency with latitude at 1030 UT is shown in Figure 5. It is seen that the resonant frequency decreases smoothly with increasing latitude, although power spectra show that power at the resonant frequency is low equatorward of ~ 70°. Parameters of the resonance were estimated as described by Menk *et al.* (2000). The resonance width in the ionosphere is of order 90–150 km and the resonance Q is 1.5 to > 3.

According to Figure 5, the 1.6 mHz signal that is prominent in the time series and power spectra is associated with a field line resonance around 71–72° latitude. Further evidence of this is presented in Figure 6, which shows the variation in amplitude and phase with latitude at a frequency of 1.6 mHz (upper panels) and 8 mHz (lower panels). These parameters were determined over 12 minute intervals using complex demodulation (Beamish *et al.*, 1979). The 1.6 mHz signal H component peaks in amplitude around 72° latitude, and the phase of both H and D components undergoes a 180° reversal around 72–73° latitude. This is characteristic of a field line resonance, but cannot explain the origin of the 1.6 mHz signals at the other stations. The most likely scenario is that 1.6 mHz waves are propagating into and throughout the magnetosphere, and couple to field line resonances where the eigenfrequency is the same, but elsewhere drive forced field line oscillations.

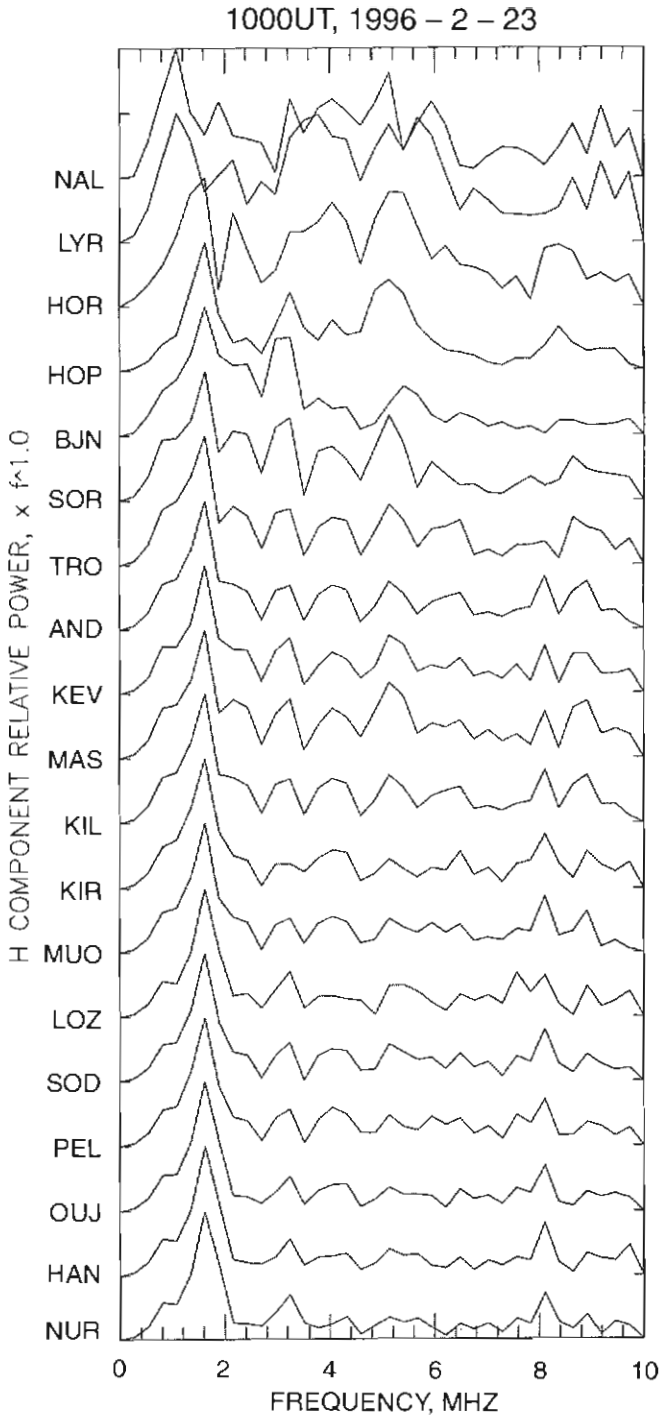


Figure 4. *Stacked normalised H component power spectra, 1000–1100 UT, 23 February 1996 weighted by* $f^{1.0}$.

The smaller, 8 mHz pulsations were present before and after the solar wind impulse. Figure 6 shows that they have maximum amplitude at the highest latitude stations, but exhibit a secondary peak in amplitude and a reversal in phase around ~66° latitude. We therefore presume the 8 mHz pulsations are also connected with inward propagating waves, that in this case stimulate field line oscillations over at least 20° in latitude, and couple to field line resonances near 66° latitude.

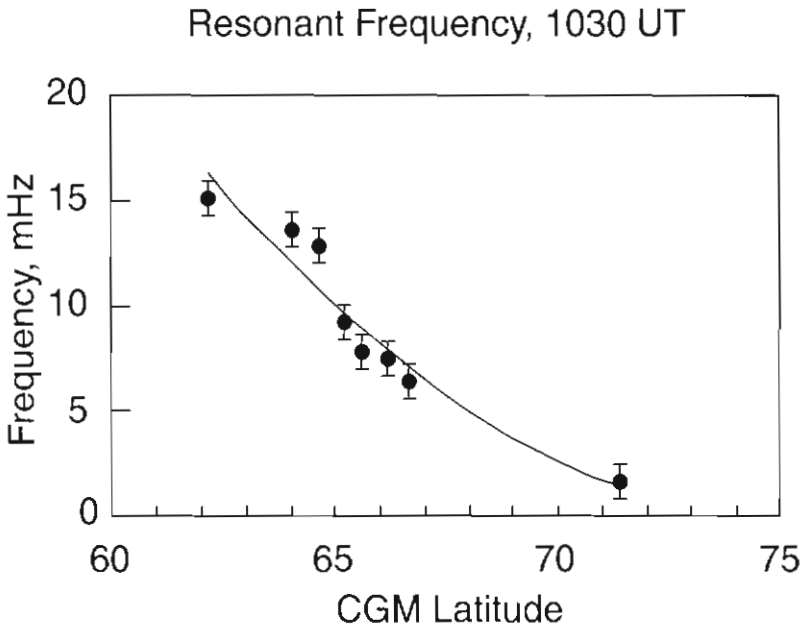


Figure 5. *Variation in field line resonance frequency determined using cross-phase measurements between adjacent IMAGE stations, 1030 UT, 23 February 1996.*

28.3.2 Radar observations

HF radar data are typically presented in the form of colour whole-day range-time parameter plots for a particular beam, where the parameters of interest are the power level of the echoes, the line-of-sight velocity, the elevation angle, and the spectral width (in m s^{-1}) of the signal. Examples may be seen at the CUTLASS homepage address given earlier. The range-time-velocity images are usually plotted over a scale of order ±1000 m s^{-1}. Backscatter regions of small velocity and spectral width are normally regarded as representing ground scatter and are therefore suppressed in the plots. In the following plots we have not suppressed the low velocity information, and in fact the plots cover a restricted velocity range of ±10 m s^{-1} or 15 m s^{-1}.

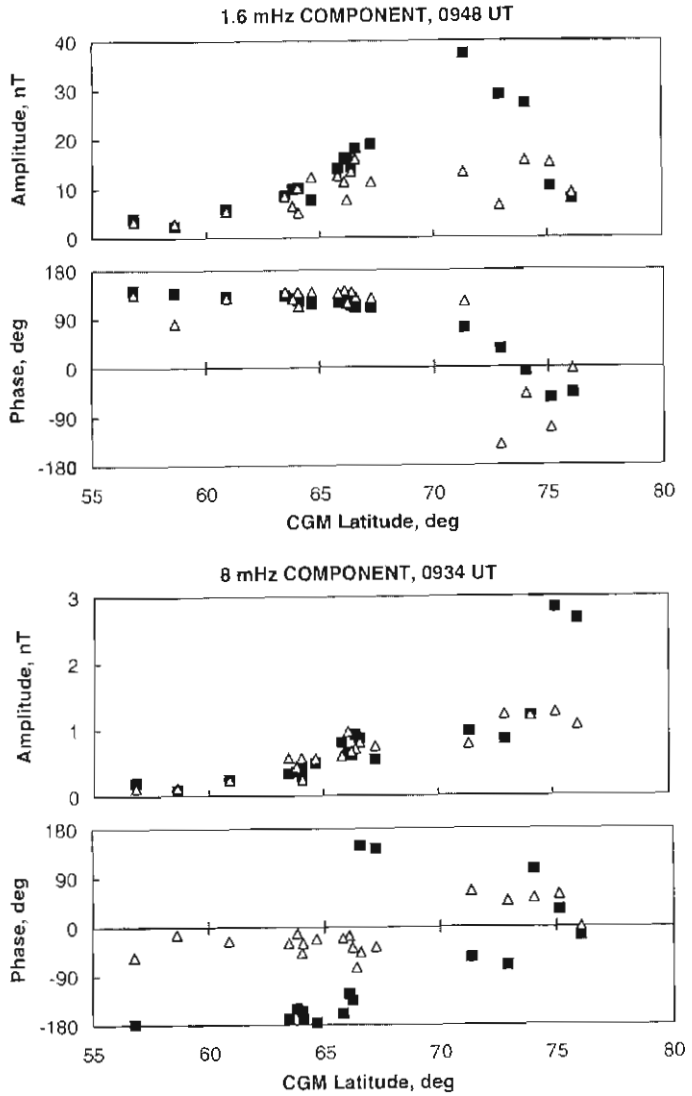


Figure 6. *Variation in pulsation amplitude and phase with latitude, at a frequency of 1.6 mHz (upper panels) and 8 mHz (lower panels), 0934 UT, 23 February 1996. Solid squares represent the H component and open triangles denote the D component.*

Figure 7(a) presents grey scale range-time-velocity plots for beams 13, 14 and 15 from the Iceland East radar, from 0900 to 1500 UT. Note that the non-standard high resolution mode did not commence until 1000 UT. Figure 7(b) shows the corresponding range-time-velocity plots for beams 5, 6 and 7 of the Finland radar, over the same time span. The Iceland East plot range axis is calibrated in 45 km range gates (commencing at 180 km), while for the Finland radar range is in magnetic latitude. The figures show many complicated features. Milan *et al.* (1999) have

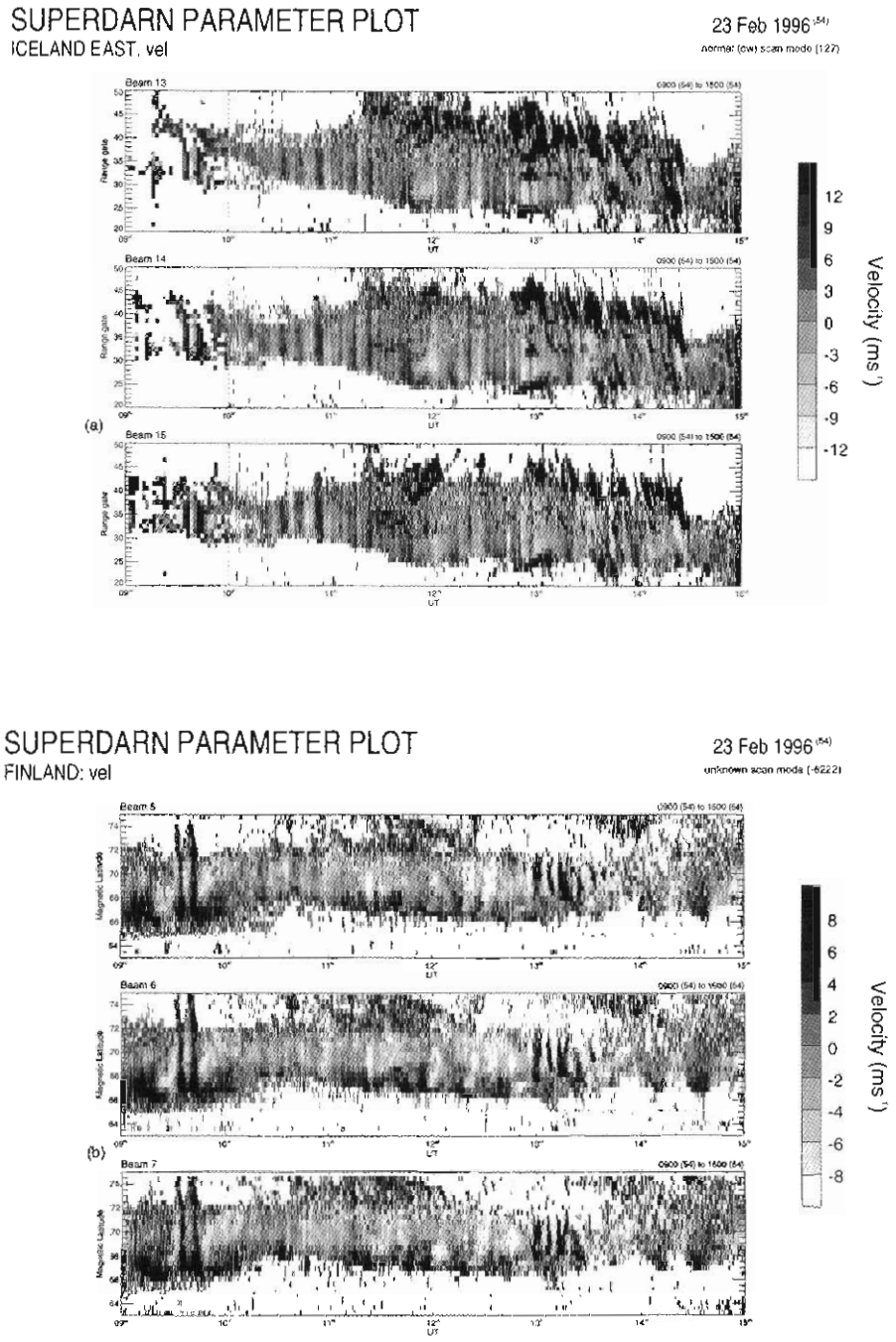


Figure 7. *Range-time-velocity plots over 0900–1500 UT, 23 February 1996 for (a) beams 13, 14 and 15 of the Iceland East radar, and (b) beams 5, 6 and 7 of the Finland radar.*

presented plots of the Finland beam five power, velocity and elevation angle for this day, and gave a detailed discussion of the features observed. Below we summarise the essential features.

For the Iceland East radar:

(i) Prominent bands of westward motion commenced around 0930 UT and extended over about 10 range gates (~ 450 km). These large scale flows showed a 10 minute periodicity and lasted some hours, gradually moving westward, toward the radar. Line-of-sight velocities were around ±10–20 m s^{-1}. Prikryl *et al.* (1998) reported similar flow features (but with much higher velocities) in HF backscatter data from the cusp region in connection with solar wind pressure pulses. They called these flow channel events (FCE's).

(ii) A more detailed stacked velocity-time-range plot for beam 15 is presented in Figure 8(a). The velocity features are essentially simultaneous across the field-of-view.

(iii) Range-time-power plots (not presented here) show that these bands were associated with enhanced power levels.

(iv) Figure 9(a) shows the spectral width in the same format as Figure 7(a), while Figure 9(b) is a plot of elevation angle for beam 14. The spectral width from the region of interest is quite small, and elevation angle relatively high. These radar returns are therefore interpreted as ground backscatter (see Milan *et al.*, 1997b).

(v) Power and phase spectra of the velocity features (see Figure 10) show a prominent peak at 1.6 mHz which has the same phase across at least 400 km in longitude.

(vi) Plots for beams 13, 14 and 15 are similar although not identical.

For the Finland radar:

(i) A prominent zone of high velocity poleward moving features around 76–80° latitude at 0900 UT moving to 72–76° latitude by 1300 UT (not shown in Figure 7(b), but see Milan *et al.*, 1999) was associated with high backscattered power and large spectral widths. This is the signature of the dayside auroral oval.

(ii) Bands of equatorward moving features commenced at ~ 0930 UT, extending from 66° to >70° in latitude. Line-of-sight velocities were of order ±20 m s^{-1}.

(iii) Figure 8(b) shows the velocity structures in more detail, for a number of range gates. The features obviously have very similar shape over 4–8 gates, i.e. latitudinal extent of order 180–360 km.

(iv) Similar results were obtained for the spectral width, elevation angles, power and phase spectra as for the Iceland East radar.

We emphasise that the features we are examining in the HF radar data are present in the ground backscatter, and are not returns from field-aligned irregularity structures in the usual sense.

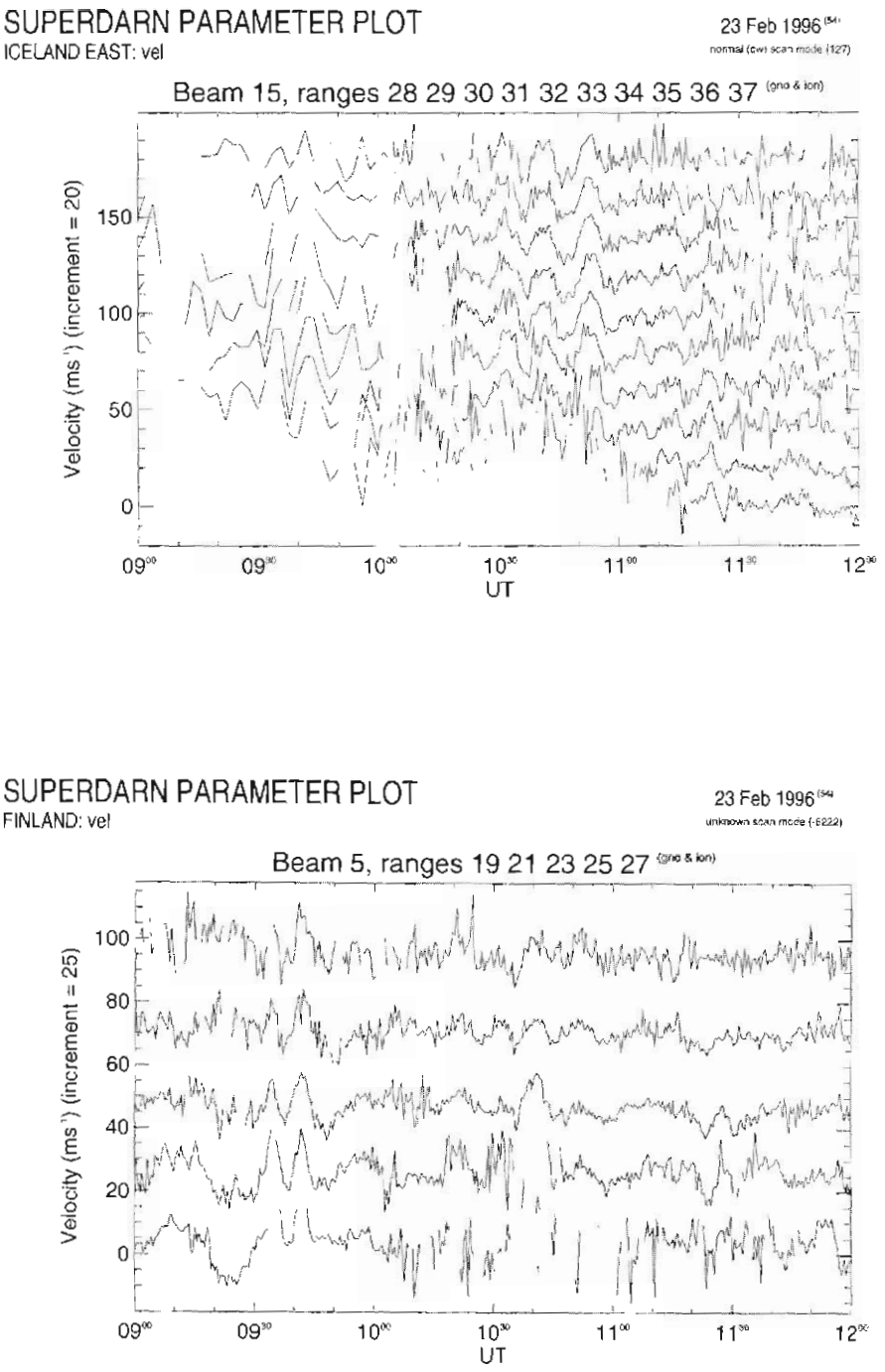
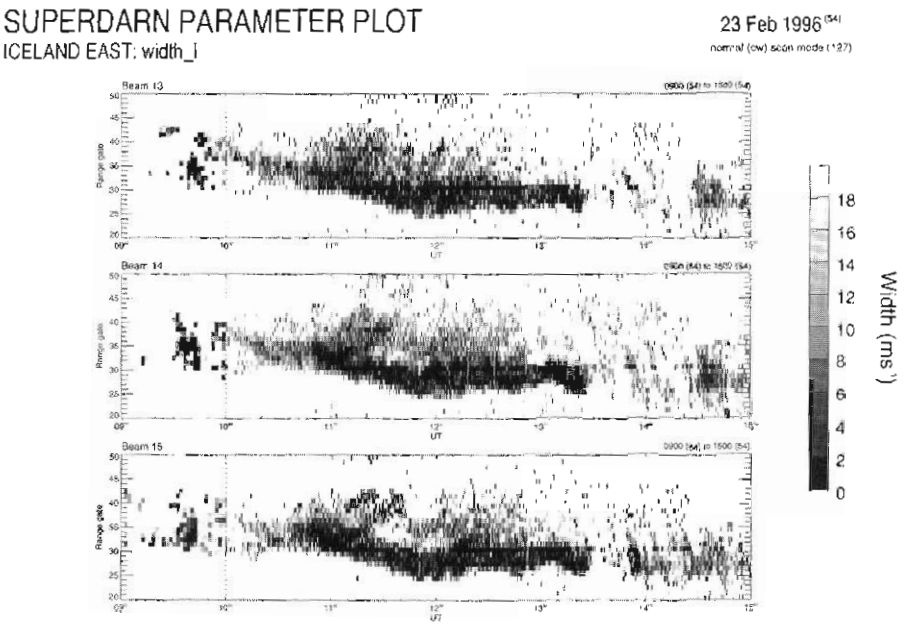


Figure 8. *Stacked velocity-time plots over 0900–1200 UT, 23 February 1996 for several range gates from (a) beam 15 of the Iceland East radar, and (b) beam 5 of the Finland radar.*



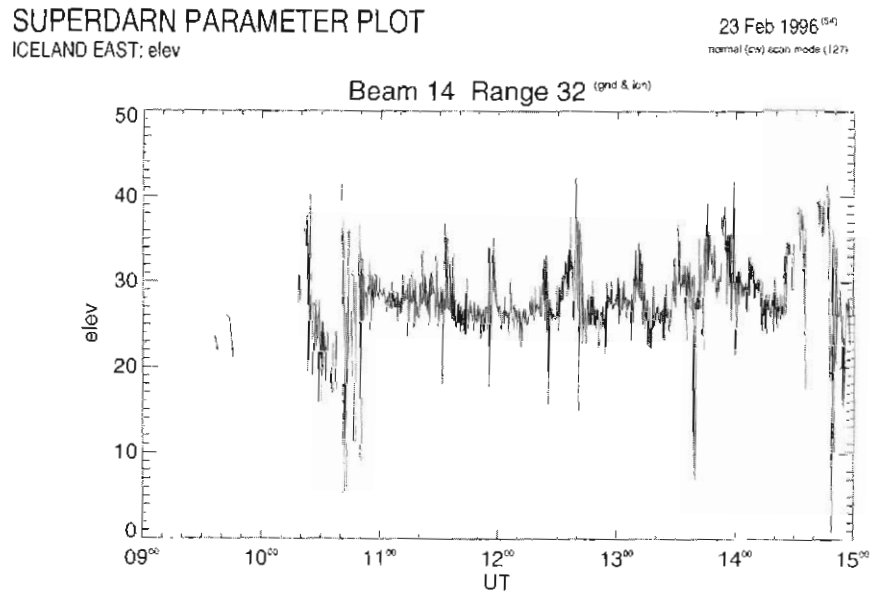


Figure 9. (a) Spectral width, and (b) elevation angle for the Iceland East radar, 0900–1500 UT, 23 February 1996.

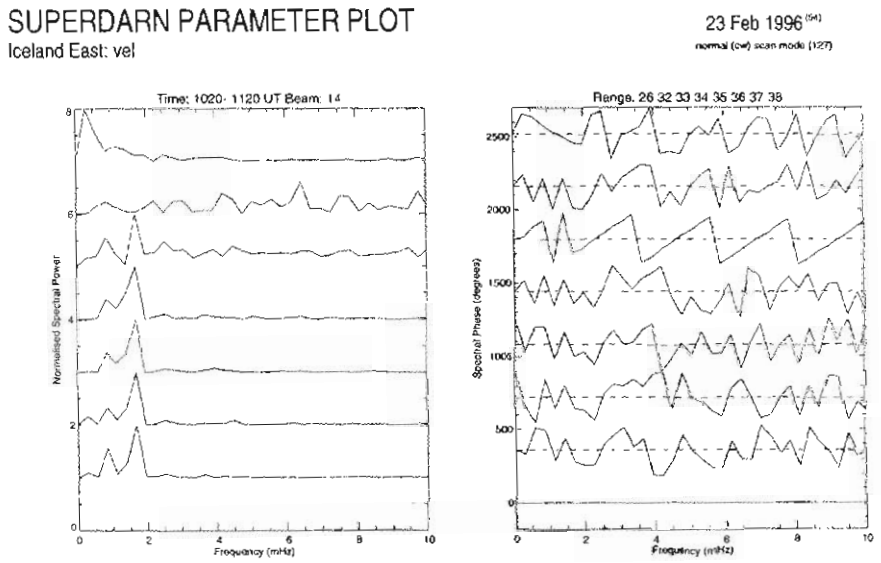


Figure 10. *Stacked spectra for selected range gates of (a) normalised power, and (b) phase, for the Iceland East radar, 1020-1120 UT, 23 February 1996. Compare with magnetometer spectra in Figure 4.*

28.3.3 Imaging riometer observations

Imaging riometer absorption maps for the interval 0920–1128 UT are presented in Figure 11. The plots represent cosmic noise absorption in dB measured each 120 s, with time moving from left to right and down the page. Geographic north is to the top and geographic west to the left. The images clearly show two absorption features to the north and just east of centre from 0920 UT until 0930 UT. The general pattern changes after 0930 UT, with lower levels of absorption across the field-of-view. Higher time resolution (15 s) images have also been examined.

The negative solar wind pressure pulse near 0930 UT is associated with a decrease in cosmic noise absorption, indicated by the decrease in intensity of the two prominent absorption patches. These patches reappear around 1000 UT, when B_y (IMF) goes negative, and again 12 minutes later, when B_y becomes positive again.

28.4 Discussion

The magnetometer observations show that the negative pressure impulse heralded the appearance of 1.6 mHz pulsations over a wide range of latitudes. Latitude independent 8 mHz pulsations were also recorded across the station array, before and after the impulse, but with smaller amplitude that the 1.6 mHz oscillations. The most likely explanation

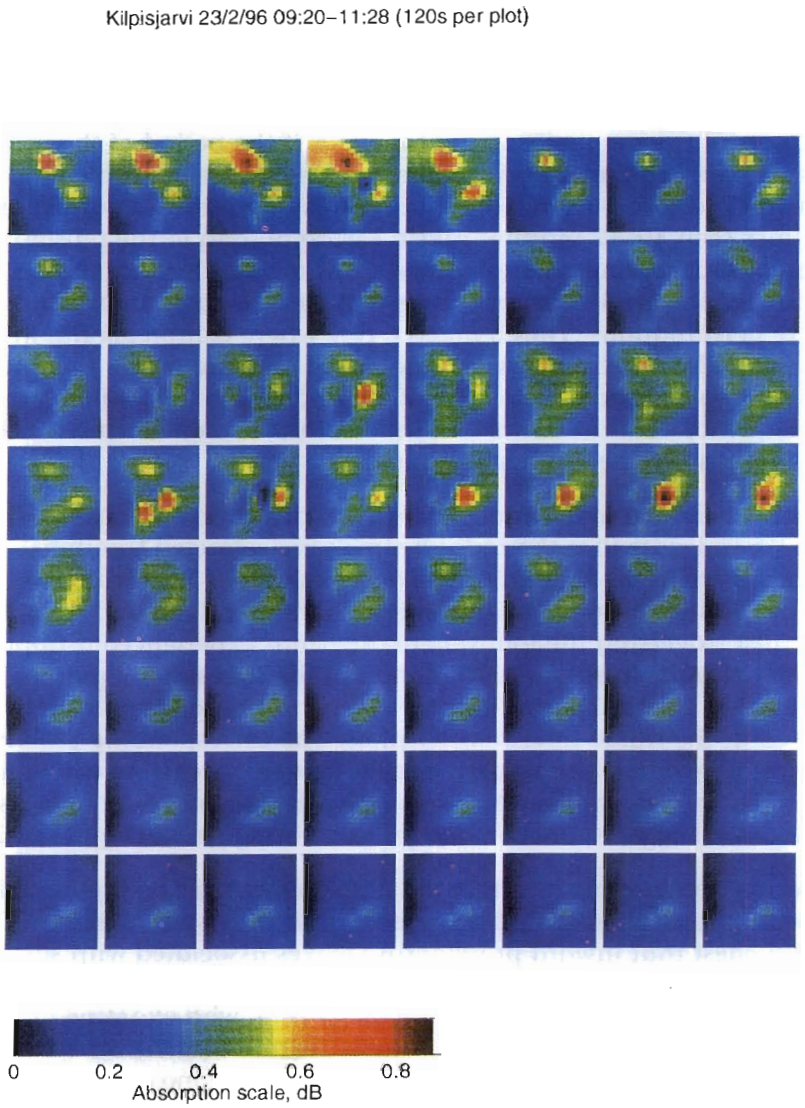


Figure 11. *Imaging riometer absorption maps for the interval 0920–1128 UT, 23 February 1996, with time moving from left to right then down the page. (See text for details.)*

of these observations is that fast mode waves were propagating into and through the magnetosphere, driving forced field line oscillations. Hasegawa *et al.* (1983) showed that such fast mode waves may also couple to resonant transverse oscillations, so that the period of the resultant pulsations depends on the latitude of the observation site. Although the 1.6 mHz pulsations dominated the time series, the cross-phase analysis summarised in Figure 5 demonstrates that latitude dependent field line resonances were also observed.

It is not clear whether the pulsations were associated with global resonances of the magnetospheric cavity, or if the period of the incoming waves is simply the rate at which the magnetopause was driven by the solar wind pressure pulses. Note that on this day pressure pulses were sometimes seen 10–12 minutes apart, corresponding to the pulsation frequency. Our magnetometer observations are similar to those reported by Prikryl *et al.* (1998), also in connection with periodic solar wind pressure perturbations. They studied a day where compressional MHD waves in the IMF associated with solar wind pressure pulses exhibited spectral peaks around 0.8, 1.3, 1.7, 2.0, 2.4 and 2.9 mHz and stimulated multifrequency compressional oscillations in the magnetosphere. These coupled to discrete field line resonances.

The most likely explanation of the low velocity bands we observed from both radars is that the radar beams were refracted through the ionosphere to the ground, then some power was backscattered from the ground through the ionosphere to the receiver, and each time whilst traversing the ionosphere the beam experienced a Doppler shift, for example in response to motion of the ionospheric plasma (e.g. Milan *et al.*, 1997b). In this case the actual range is half that indicated in the range-time-velocity plots. The corresponding regions in the ionosphere where the velocity oscillations are occurring are indicated by the solid shading in Figure 1. This is the first report of such backscatter features, and suggests a new mode for examining small Doppler shifts.

The Doppler oscillations commenced at the same time as the negative solar wind pressure pulse and occurred over a wide spatial extent. These results suggest that inward propagating waves associated with solar wind pressure pulses can drive motions of the ionospheric plasma. A similar conclusion was reached by Yeoman *et al.* (1997), who examined transient pulsations in the ionosphere. The notable feature of their work was that those pulsations were only observable because the ionosphere had been stimulated by an RF heater.

The present results demonstrate that small amplitude ionospheric oscillations are likely to be associated with impulse events even when their signature is not apparent in conventional ionospheric HF backscatter plots. For example, while Prikryl *et al.* (1998) reported FCE's with similar appearance to our velocity features in the F-region cusp, we have observed these features specifically at lower latitudes.

Based on the work of others (e.g. Sibeck, 1990; Sibeck and Samsonov, 1994) we may expect solar wind impulses to be associated with quasi-periodic enhancements in cosmic noise absorption measured by the IRIS experiment. Instead, we found the cosmic noise absorption was reduced by the initial impulse. This suggests that the negative pressure pulse resulted in a magnetospheric expansion, decreasing the energetic particle flux density. Subsequent enhancements and motions of absorption features are associated with variations in the B_y component of the IMF. The convection flow speeds we observed in the HF radar data, of order ± 10 m s^{-1}, would not result in appreciable motion of the absorption features.

28.5 Acknowledgments

IMAGE data were obtained through L. Häkkinen at the Finnish Meteorological Institute (FMI), SAMNET data from D. Milling at the University of York, and IRIS data through S. Marple at Lancaster University. The IMAGE PI is Ari Viljanen at FMI. WIND data were obtained from CDAWeb and provided by K. Ogilvie (SWE) and R. Lepping (MFI) at NASA GSFC. This work was supported by the Australian Research Council and the Particle Physics and Astronomy Research Council.

References

Beamish, D., Hanson, H.W. and Webb, D.C. (1979). Complex demodulation applied to Pi2 geomagnetic pulsations. *Geophysical Journal of the Royal Astronomical Society*, 58: 471–493.

Browne, S., Hargreaves, J.K. and Honary, B. (1995). An imaging riometer for ionospheric studies. *Electronics and Communication Engineering Journal*, 7: 209–217.

Farrugia, C.J., Freeman, M.P., Cowley, S.W.H., Southwood, D.J., Lockwood, M. and Etemadi, A. (1989). Pressure-driven magnetopause motions and attendant response on the ground. *Planetary and Space Science*, 37: 589–607.

Fejer, B.G. and Kelley, M.C. (1980). Ionospheric irregularities. *Reviews of Geophysics*, 18: 401–454.

Greenwald, R.A., Baker, K.B., Hutchins, R.A. and Hanuise, C. (1985). An HF phased-array radar for studying small-scale structure in the high-latitude ionosphere. *Radio Science*, 20: 63–79.

Greenwald, R.A., Baker, K.B., Dudeney, J.R., Pinnock, M., Jones, T.B., Thomas, E.C., Villain, J.-P., Cerisier, J.-C., Senior, C., Hanuise, C., Hunsucker, R.D., Sofko, G., Koehler, J., Nielsen, E., Pellinan, R., Walker, A.D.M., Sato, N. and Yamagishi, H. (1995). DARN/SuperDARN: a global view of the dynamics of high-altitude convection. *Space Science Reviews*, 71: 761–796.

Hasegawa, A., Tsui, K.H. and Assis, A.S. (1983). A theory of long period magnetic pulsations, 3, local field line oscillations. *Geophysical Research Letters*, 10: 765–768.

Howard, T.A., Menk, F.W. and Morris, R.J. (2001). Conjugacy of Pc3-4 waves in the high latitude magnetosphere. In: Morris, R.J. and Wilkinson, P.J. (Eds.). *ANARE Reports 146*. Australian Antarctic Division. Pp. 341–353.

Lemaire, J. (1977). Impulsive penetration of filamentary plasma elements into the magnetosphere of the Earth and Jupiter. *Planetary and Space Science*, 25: 887–890.

Lühr, H. (1994). The IMAGE magnetometer network. *STEP International Newsletter*, 4: 4–6.

Lühr, H., Aylward, A., Bucher, S.C., Pajunpää, A., Pajunpää, K., Holmbee, T. and Zalewski, S.M. (1998). Westward moving dynamic substorm features observed with the IMAGE magnetometer network and other ground-based instruments. *Annales Geophysicae*, 16: 425–440.

Lui, A.T.Y. and Sibeck, D.G. (1991). Dayside auroral activities and their implications for impulsive entry processes in the dayside magnetosphere. *Journal of Atmospheric and Terrestrial Physics*, 53: 219–229.

Menk, F.W., Waters, C.L. and Fraser, B.J. (2000). Field line resonances and waveguide modes at low latitudes; 1. observations. *Journal of Geophysical Research*, 105: 7747–7762.

Milan, S.E., Jones, T.B., Robinson, T.R., Thomas, E.C. and Yeoman, T.K. (1997a). Interferometric evidence for the observation of ground backscatter originating behind the CUTLASS coherent HF radars. *Annales Geophysicae*, 15: 29–39.

Milan, S.E., Yeoman, T.K., Lester, M., Thomas, E.C. and Jones, T.B. (1997b). Initial backscatter occurrence statistics from the CUTLASS HF radars. *Annales Geophysicae*, 15: 703–718.

Milan, S.E., Yeoman, T.K. and Lester, M. (1999). The cusp region as a hard target for coherent HF radars. *Annales Geophysicae* (in press).

Newell, P.T. and Meng, C.-I. (1988). The cusp and cleft/boundary layer: low-altitude identifications and statistical local time variation. *Journal of Geophysical Research*, 93: 14549–14556.

Newell, P.T. and Meng, C.-I. (1992). Mapping the dayside ionosphere to the magnetosphere according to particle precipitation characteristics. *Geophysical Research Letters*, 19: 609–612.

Newell, P.T., Meng, C.-I., Sibeck, D.G. and Lepping, R. (1989). Some low altitude cusp dependencies on the interplanetary magnetic field. *Journal of Geophysical Research*, 94: 8921–8927.

Newell, P.T., Wing, S., Meng, C.-I. and Sigilleto, V. (1991). The auroral oval position, structure, and intensity of precipitation from 1984 onwards: an automated online data base. *Journal of Geophysical Research*, 96: 5877–5882.

Parkhomov, V.A., Mishin, V.V. and Borovik, L.V. (1998). Long-period geomagnetic pulsations caused by the solar wind negative pressure impulse on 22 March 1979 (CDAW-6). *Annales Geophysicae*, 16: 134–139.

Potemra, T.A., Lühr, H., Zanetti, L.J., Takahashi, K., Erlandson, R.E., Marklund, G.T., Block, L.P., Blomberg, L.G. and Lepping, R.P. (1989). Multisatellite and ground-based observations of transient ULF waves. *Journal of Geophysical Research*, 94: 2543–2554.

Prikryl, P., Greenwald, R.A., Sofko, G.J., Villain, J.P., Ziesolleck, C.W.S. and Friis-Christensen, E. (1998). Solar-wind driven pulsed magnetic reconnection at the dayside magnetopause, Pc5 compressional oscillations, and field line resonances. *Journal of Geophysical Research*, 103: 17307–17322.

Pu, Z.-Y. and Kivelson, M.G. (1983). Kelvin-Helmholtz instability at the magnetopause: solutions for compressible plasmas. *Journal of Geophysical Research*, 88: 841–852.

Russell, C.T. and Elphic, R.C. (1979). ISEE observations of flux transfer events at the dayside magnetopause. *Geophysical Research Letters*, 6: 33–36.

Saunders, M.A., Russell, C.T. and Sckopke, N. (1984). Flux transfer events: scale size and interior structure. *Geophysical Research Letters*, 11: 131–134.

Sibeck, D.G. (1990). A model for the transient magnetospheric response to sudden solar wind dynamic pressure variations. *Journal of Geophysical Research*, 95: 3755–3771.

Sibeck, D.G. and Samsonov, S.N. (1994). Transient CNA events and magnetospheric compressions. In: Baker, D.N., Papitashvili, V.O. and Teague, M.J. (Eds.). *COSPAR Colloquia 5*. COSPAR. Pp. 323–326.

Takahashi, K., Kistler, L.M., Potemra, T.A., McEntire, R.W. and Zanetti, L.J. (1988). Magnetospheric ULF waves observed during the major magnetospheric compression of November 1, 1984. *Journal of Geophysical Research*, 93: 14369–14382.

Tsunoda, R.T. (1988). High-latitude F region irregularities: a review and synthesis. *Reviews of Geophysics*, 26: 719–760.

Villain, J.P., Caudal, G. and Hanuise, C. (1985). A SAFARI-EISCAT comparison between the velocity of F region small-scale irregularities and the ion drift. *Journal of Geophysical Research*, 90: 8433–8443.

Warnecke, J., Lühr, H. and Takahashi, K. (1990). Observational features of field line resonances excited by solar wind pressure variations on 4 September 1984. *Planetary and Space Science*, 38: 1517–1531.

Waters, C.L., Samson, J.C. and Donovan, E.F. (1995). The temporal variation of the frequency of high latitude field line resonances. *Journal of Geophysical Research*, 100: 7987–7996.

Yeoman, T.K., Milling, D.K. and Orr, D. (1990). Pi2 polarization patterns on the U.K. Sub-Auroral Magnetometer Network (SAMNET). *Planetary and Space Science*, 38: 589–602.

Yeoman, T.K., Wright, D.M., Robinson, T.R., Davies, J.A. and Rietveld, M. (1997). High spatial and temporal resolution observations of an impulse-driven field line resonance in radar backscatter data artificially generated with the Tromsø heater. *Annales Geophysicae*, 16: 634–644.

29. THE FEDSAT NEWMAG MAGNETIC FIELD EXPERIMENT

B.J. Fraser[1], C.T. Russell[2], J.D. Means[2], F.W. Menk[1] and C.L. Waters[1]

[1] CRC for Satellite Systems
Department of Physics
University of Newcastle
NSW 2308 Australia
(email: bjfph@cc.newcastle.edu.au)

[2] IGPP/UCLA
Los Angeles
CA 90024
USA

Abstract

Late in the year 2001 Australia will launch FedSat, a research microsatellite with a complement of four scientific and engineering payloads. It will be placed in a near-circular Sun synchronous 1030 LT polar orbit at an inclination of 98.7° and altitude 800 km. The primary scientific experiment is a fluxgate magnetometer which will observe field aligned current structures above the auroral zones and other current systems and ultra-low frequency (ULF) plasma waves in the upper ionosphere. A major objective is to study the ionospheric signatures of regular and impulsive ULF waves propagating from the magnetosphere along geomagnetic field lines. An on-board GPS receiver will monitor the vertical and oblique total electron content (TEC) up to the GPS satellites at an altitude of 20 000 km. This paper concentrates on the research to be undertaken using the fluxgate magnetometer and describes the instrumentation.

29.1 Introduction

In 1997 the Cooperative Research Centre for Satellite Systems (CRCSS) was established under the Australian Government's CRC Program. The CRC program fosters research collaboration between Australian Universities, government research organisations and industry on a seven year cycle. There are more than sixty CRC's currently active in engineering, science, medicine, and technology. The specific mission of the CRCSS is to deliver a new sustainable advantage for Australian industries and government agencies, including universities, involved in services based on the applications of future generations of small satellites.

The primary project of the CRCSS is to launch a low Earth orbiting microsatellite, FedSat into an 800 km circular polar orbit Sun synchronous at 1030 LT with an inclination of 98.7°. The launch, a piggyback ride on a Japanese H-2A rocket with ADEOS-2, will take place towards the end of the year 2001. FedSat will become operational at the time of celebrations of Australia's Centenary of Federation. FedSat will have a nominal mass of 50 kg with the approximate dimensions of a 50 cm cube and the addition of a 2.5 m deployable boom. The proposed satellite is illustrated in Figure 1. The satellite platform is under construction at Space Innovations Limited (SIL), a United Kingdom company, and is based on their MicroSILTM microsatellite bus. The attitude control system employs a three axis stabilisation system and is nadir pointing.

FedSat is not a mission dedicated to a single scientific or engineering goal. Instead, it will carry a range of experimental research payloads which are based on the engineering and scientific research activities of the participants in the CRCSS. The aim is to allow experience to be gained in all areas. Payloads include a fluxgate magnetometer; a UHF data transfer system; a Ka-band communications system; precise orbit determination and TEC measurements using GPS; and a high performance computer experiment. These CRCSS activities are focussed under four programs, the Space Science Program (University of Newcastle; La Trobe University), the Communications Program (University of South Australia; University of Technology; DSTO), the Satellite Systems Program including navigation and the high

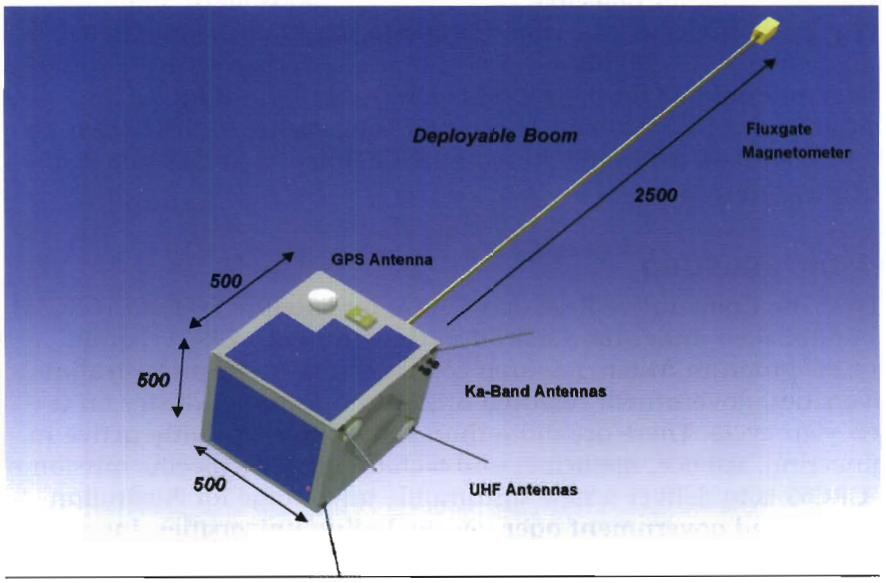


Figure 1. *The proposed FedSat microsatellite in orbit.*

performance computer experiment (HPCE) (Queensland University of Technology) and the Satellite Engineering (CSIRO; Auspace Ltd.; Vipac Engineers and Scientists). There is also an Education Program in the CRCSS.

This paper is primarily concerned with the Space Science Program and in particular the magnetometer experiment. Further details on the FedSat platform are given in Barrington-Brown *et al.* (1998) and Graham (1999).

29.2 The NEWMAG science project

The overall objectives of the experiment are to:

(i) Conduct basic research on the structure and dynamics of the ionosphere and exosphere/inner plasmasphere using magnetic field observations and to apply the results to diagnostics and where relevant space weather modelling. By including the propagation properties of GPS signals the effects of upper atmosphere disturbances on communication systems may also be studied.

(ii) Study the dynamics of field-aligned currents and filamentary structures in the auroral zones, and the equatorial electrojet current system.

(iii) Study oscillating wave fields and their variability in the ionosphere and exosphere, including lower ELF band ion cyclotron waves and ULF (ultra-low frequency) hydromagnetic waves in conjunction with Antarctic and middle/low latitude ground station observations.

(iv) Contribute to the monitoring and mapping of the geomagnetic field in the Australian region.

The instruments that contribute to these objectives are the NEWMAG magnetometer experiment and the ionospheric TEC analysis using GPS.

29.2.1 The near-Earth plasma environment

The ionosphere and exosphere represents a transition between the highly ionised plasma dominated magnetosphere and the lower altitude neutral atmosphere. The topside ionosphere is an important boundary region that exerts an important influence on the evolution of the planetary atmosphere. In the regions above the auroral zones in the northern and southern hemispheres the topside ionosphere is continuously being replenished with plasma at the rate of several kg s^{-1} (Chappell *et al.*, 1987). This plasma outflow is due to heating and energisation processes associated with downward electron acceleration and upward and transverse acceleration of ions. Although FedSat will orbit at 800 km, below the 1000–10 000 km altitude of the acceleration region, it will pass through the region connecting currents and propagating wave energy to the ionosphere.

The behaviour of a magnetised plasma can be described in terms of the configuration of the magnetic field or the currents that are flowing. In examining the coupling of the ionosphere to the magnetosphere it is traditional to describe these in terms of field-aligned current systems. The

region 1 and region 2 current systems (Iijima and Potemra, 1976) couple the stresses applied to the magnetospheric plasma by the solar wind and internal convection of the plasma. A schematic of the magnetospheric current systems is shown in Figure 2.

29.2.2 Field-aligned currents

Figure 3 shows the magnetic fields measured by the FAST (Fast Auroral SnapshoT Explorer) spacecraft magnetometer with the Earth's field removed, on a pass through the dayside auroral oval and across the polar cap. The antisunward twists in the field at about 65° latitude are the characteristic signatures of field-aligned currents. The currents flowing near midnight around the onset of substorms, the so-called substorm current wedge (McPherron, 1979), are believed to arise from the rapid reconfiguration of the tail when magnetic reconnection in the tail begins. There is not consensus on the mechanisms by which these stresses arise and this particular problem will be addressed through the FedSat fluxgate magnetometer measurements. Additional studies to be undertaken include identifying polar cusp magnetic signatures, and relating low altitude observations with higher altitude spacecraft including IMAGE and CLUSTER using field line tracing and simulations. Also important will be correlation with ground observations made at the Australian Antarctic stations of Davis, Casey and Mawson, and the USA Automatic Geophysical Observatories located on the southern polar cap.

Ultra-low frequency (ULF) waves, naturally occurring plasma waves in the 1 mHz–5 Hz band, are seen in all regions of the magnetosphere

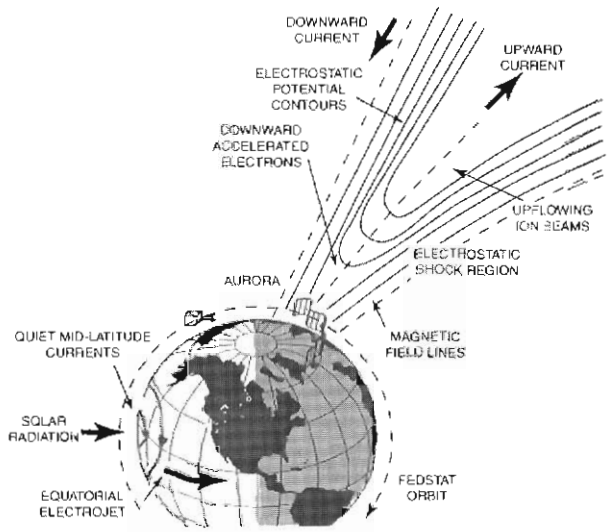


Figure 2. *Magnetospheric current systems.*

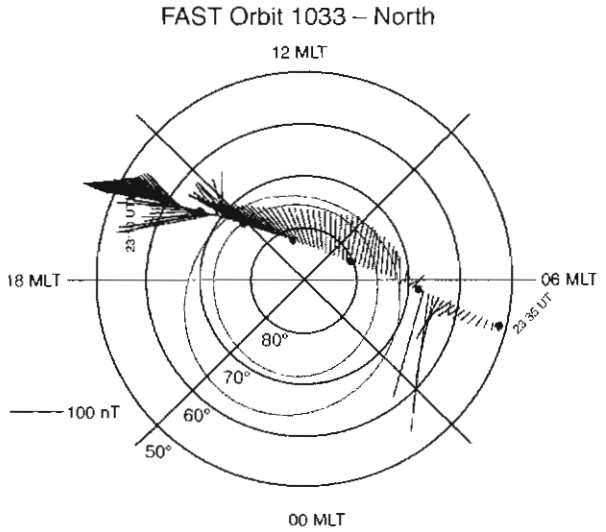


Figure 3. *Residual magnetic fields with the model IGRF removed on a FAST pass over the northern polar cap. The anti-sunward residuals near 65° are the signatures of region 1 and region 2 currents. The FedSat magnetometer will be capable of improving the reliability of these results.*

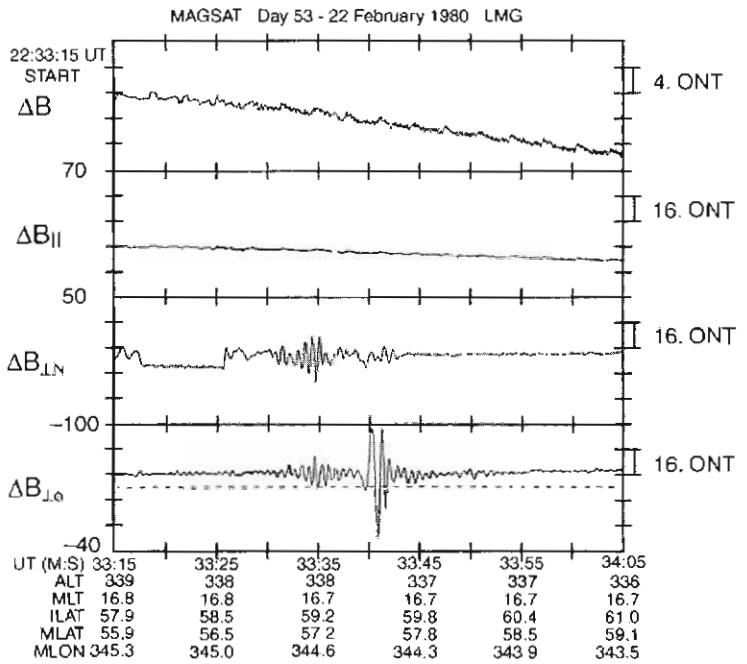


Figure 4. *An example of a Pc1 ion cyclotron wave signature seen by MAGSAT. The impulsive wave frequency matches that of a sustained Pc1 event seen on the ground (Iyemori and Hayashi, 1989).*

from high to equatorial latitudes (Anderson *et al.*, 1990, 1992). Many types of ULF waves are also transmitted from the magnetosphere through the ionosphere to the ground. Since the ionosphere acts as a 'short circuit' or sink for many magnetospheric phenomena, ground-based and ionospheric signatures play an important role in understanding the interactions with solar energy sources. Interactions that will be observed by the FedSat magnetometer and Antarctic and other ground stations include flux transfer events (FTE's) which are signatures of transient dayside reconnection (Russell and Elphic, 1978) and solar wind pressure and other variations. These signatures can also be seen at middle and low latitudes.

The more regular Pc3-5 ULF waves in the 1–100 mHz band are generally observed signatures of resonances in the magnetosphere, while the higher frequency Pc1-2 (0.1–5 Hz) waves result from wave-particle interactions in the magnetosphere, or low altitude resonances. At Pc5 frequencies using VIKING data Potemra *et al.* (1988) have observed resonant geomagnetic field line oscillations on the same field lines that guide region 2 currents in the morning sector. We will be looking for lower altitude signatures using FedSat. With these longer period waves it is important to separate the spatial properties associated with spacecraft motion from the temporal properties.

In the higher frequency Pc1-2 ULF band wave-particle interaction in the equatorial region of the middle and outer magnetosphere generates electromagnetic ion cyclotron (EMIC) wave packets which propagate down field lines to the ionosphere (Fraser, 1985). Impulsive magnetic signatures associated with these have been observed on the MAGSAT satellite at an altitude of 400 km by Iyemori and Hayashi (1989) and associated with ground observations. An example is shown in Figure 4. These will also be seen by FedSat in its polar orbit. When the electromagnetic ion cyclotron (EMIC) wave packets propagating down the field line reach the ionosphere some of their energy is reflected and some is mode-converted from the left-hand polarised ion cyclotron wave into the right-hand polarised isotropic fast mode wave. It is well known that two distinct waveguides exist in the ionosphere. One is centred on the F_2 region electron density peak with a width ~ 300 km and supports the horizontal propagation of wave energy at f ~ 1–5 Hz over many thousands of kilometers, especially at night when attenuation is low (Fraser, 1975). A second and much wider resonance region exhibiting harmonics in the 0.3–10 Hz band, the so called sub-protonospheric cavity (Trakhtengertz and Feldstein, 1987), has upper and lower boundaries at the Alfvén velocity gradient at ~ 3000 km and the E-region, respectively. It is expected that waves in both waveguides would be observed by FedSat as it traverses the middle and high-latitudes. Because of their small-scale structure these propagating waves may be used as tracers in diagnostics to improve our knowledge of the ionosphere and exosphere/ inner plasmasphere.

29.2.3 Lower-band ELF waves

In the high-latitude auroral and cusp regions there are many propagating wave modes which are generated in boundary or acceleration regions. The properties of these waves are generally controlled by the plasma characteristic frequencies, primarily the H^+ cyclotron frequency and the corresponding O^+ and He^+ heavy ion frequencies. At high-latitudes and lower exosphere altitudes the H^+ cyclotron frequency is typically ~ 700 Hz, the He^+ cyclotron frequency ~ 200 Hz, and the O^+ cyclotron frequency ~ 40 Hz. With an upper limit on the FedSat fluxgate magnetometer bandwidth of 50 Hz it will only be possible to observe electromagnetic waves near and below the O^+ cyclotron frequency. Plasma turbulence provides a significant contribution to the electromagnetic spectrum at these frequencies. There is a rich spectrum of waves below the H^+ cyclotron frequency and examples of these have been given by Erlandson (1989).

In addition to the waves described above there is also a tropospheric contribution to the ELF wave flux in the ionosphere. A lightning strike in the atmosphere generates a broad band electro-magnetic pulse that is absorbed at the local proton cyclotron frequency as it propagates upwards. The usually right-hand polarised wave is converted to a left-hand wave at the H^+ crossover frequency in the multicomponent ionospheric plasma. These are ion cyclotron whistlers (Gurnett *et al.*, 1965). Similar resonances with He^+ and O^+ ions have also been observed, with the latter occurring in the FedSat magnetometer bandwidth.

The recent discovery of lightning discharges between the tops of clouds and the lower ionosphere, commonly referred to as red sprites and blue jets (Sentman *et al.*, 1995) has caused a rethinking of our understanding of atmospheric electricity. The red sprites also radiate at VLF frequencies. It is not known whether a further type of optical emission seen in the lower ionosphere referred to as elves (Fukunishi *et al.*, 1996), has an ELF electromagnetic signature. It is important to look for coupling processes at all frequencies which may transfer energy from the troposphere to the ionosphere. FedSat will pass over the south-east Asian thunderstorm region between November and March.

29.2.4 Equatorial electrojet

Currents in the ionosphere above the equatorial region have not been studied to any great extent. Although this current system is more difficult to measure by satellite than the auroral system, it is equally dynamic with upflowing plasma and plasma bubbles playing an important role (Woodman and LaHoz, 1976). This region is of great importance as scintillation and irregularities can affect modern day communications.

29.2.5 Space weather forecasting

It is now known that solar induced variability in the near-Earth space environment can affect the performance of technological systems. High energy magnetic cloud events outbursting from the Sun can impinge on the Earth and destroy satellites, disrupt power grid systems, HF

radio communications and GPS navigation systems. These disruptions will maximise at the peak of the sunspot cycle at the time when FedSat is expected to be launched. Continuous multipoint spacecraft measurements of the global magnetic field and plasma environment are required for the early warning of impending events and inclusion in space weather prediction models currently being developed. FedSat will make an important contribution here. The NEWMAG experiment will provide near real-time data to the Space Environment Centre of NOAA in Boulder Colorado and the IPS Radio and Space Services in Sydney for forecasting purposes.

29.2.6 The geomagnetic field in the Australian region

The Australian Geological Survey Organisation (AGSO) provides reference magnetic field data to the geophysical exploration and mining communities for mineral exploration. Since the geomagnetic field is continuously changing, this information must be updated every few years and is usually undertaken through expensive ground magnetometer surveys (Barton, 1997). Spacecraft observations of sufficient accuracy have not been available for a number of years since MAGSAT in 1979–80. This has recently been improved by the launch of Oersted and SunSat, which use star imagers to precisely determine magnetometer sensor head attitude. In a recent development a star imager will be flown on the FedSat platform, thereby providing a greatly improved knowledge of attitude, hopefully down to a few tens of arc seconds. It is hoped that data from the highly stable and accurately calibrated FedSat fluxgate magnetometer will be used in conjunction with data from dedicated magnetic field satellites, including Oersted and SunSat, and CHAMP, and SAC-C in the future, to help upgrade magnetic field measurements over Australia and other regions.

29.2.7 Monitoring the ionosphere and plasmasphere

Data from this project complements the NEWMAG data. The aim of this project is to use GPS data and advanced processing to further our understanding of the ionosphere and plasmasphere and improve current ionospheric models. The GPS network of stations on the ground and in orbit provides a unique opportunity to make these measurements. Total electron content (TEC) measurements will be undertaken between GPS satellites at apogee (20 000 km) and the Spectrum-Astro GPS receiver on-board FedSat at 800 km altitude, thereby providing information on the exosphere inner plasmasphere. TEC occultation observations made with a number of GPS satellites will provide slant measurement which may be converted into a vertical profile. The data will also be used to provide electron density distributions over the southern oceans and Antarctica. Further details of the FedSat TEC measurements and their applications are included in this volume, Essex *et al.* (2001).

29.3 The FedSat platform

The multiple payloads identified in Section 29.1 above will be accommodated in a platform based on the Space Innovations Limited

(UK) MicroSILTM satellite bus. The subsystems are based on SIL's suite of standard on-board equipment with minor modifications. The FedSat bus, seen in Figure 5, is a two tiered shelf structure with the payloads on the top shelf and the subsystems on the bottom shelf. The subsystems are load-bearing, thereby providing considerable mass and volume savings. The interfacing is simple with the payload shelf consisting of two panels, one interfacing the subsystems and the second for payload use.

The payloads within the spacecraft are accommodated in three separate enclosures as shown in Figure 5. One contains the GPS receiver, the second the communications experiment and the third, the magnetometer analogue and digital electronics, and the Higher Performance Computer Experiment (HPCE). The magnetometer experiment requires the spacecraft design to incorporate a magnetic cleanliness program to minimise both steady state and time varying magnetic fields. The fluxgate heads are mounted at the end of a 2.5 m boom. This boom has been built by Stellenbosch University and is same as that successfully used on the South African micro-satellite SunSat. It consists of six segments with flexible interconnections and the fluxgate heads are mounted at the end. It is seen in its extended position in Figure 1.

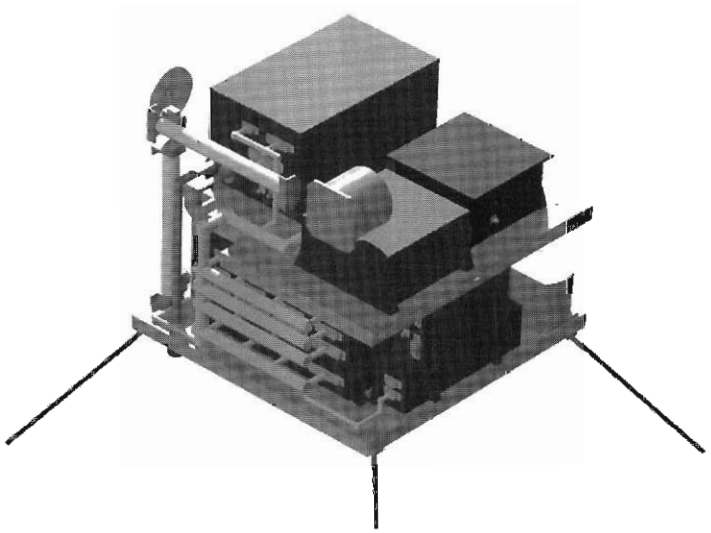


Figure 5. *A view of the basic MicroSIL platform structure.*

The attitude control system employs three rod magnetotorquers, a three axis fluxgate magnetometer, four reaction wheels (one redundant) and three two axis digital sun sensors. This is a new Dynacon (Canada) system which should provide sufficient pointing knowledge to maintain the satellite pointing better than 1° of the target direction.

A block diagram of the FedSat electrical system is illustrated in Figure 6. More details on the FedSat platform and subsystems are included in Barrington-Brown *et al.* (1998) and Graham (1999).

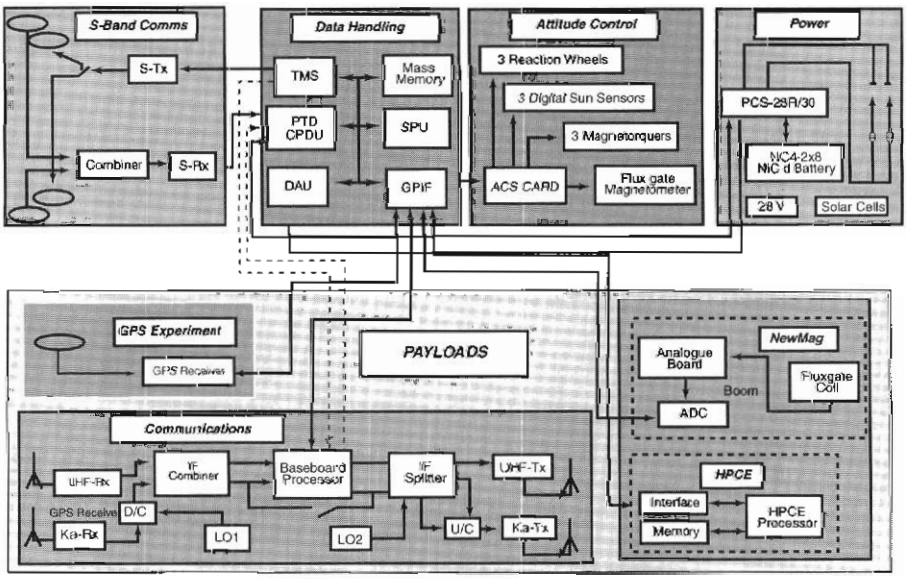


Figure 6. *The MicroSIL platform electrical block diagram.*

29.4 FedSat magnetometer instrumentation

The FedSat triaxial fluxgate magnetometer is a high precision instrument designed to measure magnetic fields at all latitudes at an 800 km altitude over a range of ±65 000 nT. The magnetometer needs to be sufficiently sensitive to measure the magnetic perturbations associated with the various current sheets and wave processes referred to in Section 29.2. In the polar regions where the geomagnetic field is strongest FedSat will measure field intensities up to approximately 30 000 nT and superimposed perturbations of the order of 10 nT. In addition, field-aligned current structures with filamentary current sheets will be traversed rapidly. Scale sizes vary from ~ 1000 km down to 5 km and possibly smaller, with corresponding signatures of several hundreds of nT down to 5 nT respectively (Weimer *et al.*, 1985). Since the spacecraft moves at ~ 8 km s^{-1} a sample rate of ten vector samples per second (V s^{-1}) is required to resolve 1 km structures. This is the limit of the FedSat fluxgate background sampling rate but the burst mode operation at 100 V s^{-1} will be capable of resolving even smaller current structures. ULF waves in the 1–100 nT range will be measured with highest

amplitudes seen at high-latitudes. Pc1 and ELF have, over the 1–50 Hz band, lower amplitudes and will be more difficult to detect and pure state vector analysis techniques will be employed (Samson, 1983).

29.4.1 Magnetometer description

The instrument is being built at the Institute of Geophysics and Planetary Physics, UCLA by C.T. Russell and has advantages of a sustained spacecraft magnetometer heritage which includes amongst others, the ISEE-1/2 AMPTE-UKS, Galileo, POLAR and FAST fluxgate magnetometers. The FedSat triaxial fluxgate magnetometer takes advantage of the POLAR (Russell *et al.*, 1995) and FAST (Elphic *et al.*, 2001) analog designs in order to minimise development costs and fabrication time. The main characteristics of the system are listed in Table 1.

29.4.2 Fluxgate sensors

The sensors are fabricated using ring core sensors built at UCLA. Each of the three sensors contains a drive winding and a sense/feedback coil surrounding a magnetically permeable core. Correctly phased currents in the drive coil saturate the ring core twice every drive cycle. External magnetic fields upset the symmetry of these saturations, resulting in the generation of even harmonic signals whose amplitude is proportional to the ambient field, in this situation, the geomagnetic field. These signals are detected by the sense coil and used in a feedback system to maintain the core at approximately zero field. The drive coils will be driven in parallel and the sensor assembly will be mounted at the end of the 2.5 m boom.

Table 1. *The NEWMAG Magnetic Field Experiment.*

Characteristics:	
Triaxial ring core fluxgate sensors:	
	Mass 420 g
	Dimensions 11.4 x 5.4 x 6.1 cm
	Power < 50 mW
Electronics:	
Analogue	
	Mass 600 g
	Power ~ 2 W
ADC/DPU	
	ADC 16 bit
	Mass 600 g
	Power ~ 1 W
Dynamic range	±65,000 nT
Digitisation	±0. nT
Sample rates:	
Internal sample rate	1000 V s^{-1}
Standard data rate	10 V s^{-1}
Burstmode/snapshot	100 V s^{-1}

V s^{-1} ≡ vector samples per second.

29.4.3 Electronics

The NEWMAG on-board electronics are mounted on two boards in a box on the payload shelf in the platform and includes the fluxgate analogue circuits, analog-to-digital converter, data processing unit (DPU) and interfacing circuits.

The analog board follows the design of the POLAR system (see Figure 6, Russell *et al.*, 1995). This circuitry provides the precisely timed signal to the ring cores at a frequency of 9 kHz. The amount of second harmonic signal in quadrature with the drive frequency is detected. The feedback circuit nulls this harmonic by providing sufficient current to keep the ring cores in zero fields. This current indicates the strength of the external field. The low pass filters provide output signals with a cutoff frequency at 1000 Hz.

The digital processing board will include an analogue-to-digital converter (ADC) and a microprocessor controller with filtering capabilities. A 16 bit Crystal or similar ADC will be used to cover the ±65 000 nT field range with an accuracy of ±1 nT. The microprocessor controls the operations of data sampling and filtering (DPU). A flow diagram of the on-board data processing is included in Figure 7. The 100 V s^{-1} (vector samples per second) data from the analog board will be multiplexed and fed into the ADC. The ADC will sample each channel at 1000 Hz and average the data down to 10 V s^{-1}. By undertaking this averaging process the accuracy of the measurements are extended to 18 bit resolution, equivalent to a ±0.2 nT digitisation window. Both the 100 V s^{-1} and 10 V s^{-1} data will be stored in the data processing unit

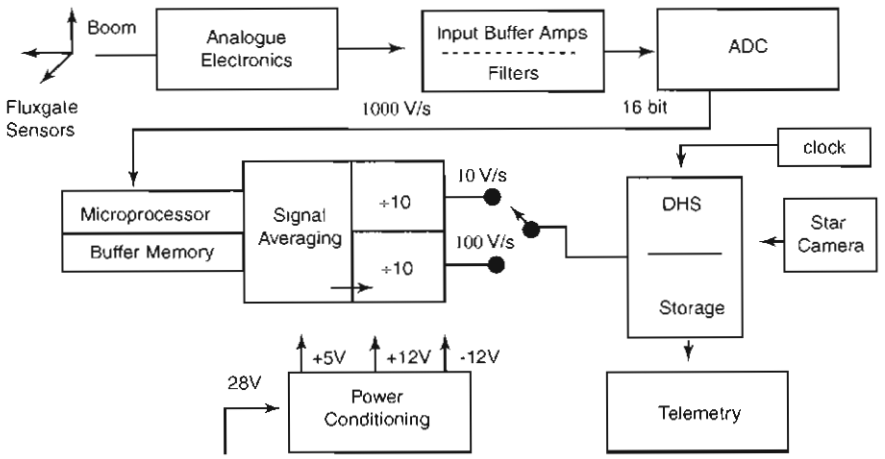


Figure 7. *The on-board analogue magnetometer and digital processing system.*

(memory/processor) and fed to the platform data handling system (DHS) periodically for storage and later transmission to ground.

FedSat has the advantage of three axis stabilisation and consequently it will not experience a spin frequency component in the geomagnetic field data. The 100 V s^{-1} data which will be recorded in burst mode on command. It is anticipated that up to 40 snapshots a day of 100–200 s duration may be taken if storage is available. The recording of a snapshot burst mode interval could be initiated by direct command, or a preset event level trigger. This recording mode is most useful in the auroral zone when FedSat is passing through filamentary current structures. The data processing unit may be capable of performing additional computations such as Fast Fourier Transformations leading to a nine element cross-spectral matrix over individual snapshots and also a sequential ordering of the amplitude of snapshots. Data compression to eight bits may also be undertaken to conserve storage.

29.5 Data analysis and archiving

The magnetometer data will be transmitted to a ground station located at the University of South Australia in Adelaide. In the first year of operation UCLA and the University of Newcastle will be responsible for developing and implementing algorithms to decode the level zero data. The routine calculation will then be taken over by the University of Newcastle. All data will be validated by comparison with models such as the well-established International Geomagnetic Reference Field (IGRF) over regions of low perturbation by external current systems (e.g. at low latitudes) where the Earth's field is well-defined. The absence of a spacecraft spin component in the magnetic data enables the coordinate transformation to be accomplished much more accurately and quickly than with a spinning spacecraft. The detection of low level signal will be enhanced using pure state multichannel time series analysis techniques (Samson, 1983). All raw and level one data will be archived at UCLA and the University of Newcastle. It will also be freely available to the international scientific community on a website and through the National Scientific Satellite Data Centre (NSSDC) in Washington DC.

In addition to the international availability of FedSat magnetic field data Australian research collaborations are being established. Satellite-ground studies of auroral processes and waves above the Australian Antarctic stations at Casey, Davis, Mawson, and Macquarie Island will be undertaken with the Australian Antarctic Division. Near real time data will be made available directly to the IPS Radio and Space Services (IPS) in Sydney. IPS is a member of a group of world-wide ionospheric and space weather warning and forecasting centres. Main geomagnetic field measurements if proven sufficiently accurate will also contribute to the research of the Australian Geological Survey Organisation (AGSO) through the mapping of the geomagnetic field above the Australian region. Ideally, these studies would be more reliable and useful if FedSat was at a lower altitude. However, the inclusion of a high precision star

imager for precise attitude determination down to tens of arc seconds will enhance this capability. It is therefore important to accurately calibrate and test the FedSat magnetometer prior to launch and on orbit.

29.6 Conclusion

A description of the NEWMAG magnetometer experiment on-board FedSat, an Australian research microsatellite to be launched late in 2001, has been presented. Included are broad details of the scientific studies to be undertaken. Rather than being a satellite mission dedicated to a specific research task it is a satellite which takes advantage of current Australian space science, communications and satellite systems research capabilities and integrates them with space engineering expertise. It will lead Australia back into space after thirty years absence while making a significant contribution to space science and engineering.

29.7 Acknowledgements

We gratefully acknowledge the efforts of many individuals within the CRC for Satellite Systems, the Institute of Geophysics UCLA, and the University of Newcastle. This work was carried out in the Cooperative Research Centre for Satellite Systems with financial support from the Commonwealth of Australia through the CRC program, and the University of Newcastle.

References

Anderson, B.J., Engebretson, M.J., Rounds, S.P., Zanetti, L.J. and Potemra, T.A. (1990). A statistical study of Pc3-5 pulsations observed by the AMPTE/CCE magnetic field experiment. 1. occurrence distributions. *Journal of Geophysical Research*, 95: 10495–10523.

Anderson, B.J., Erlandson, R.E. and Zanetti, L.J. (1992). A statistical study of Pc1-2 magnetic pulsations in the equatorial magnetosphere. 1. equatorial occurrence distributions. *Journal of Geophysical Research*, 97: 3075–3088.

Barrington-Brown, A.J., Wicks, A.N., Boland, L., Gardner, S.J. and Graham, E.C. (1998). FedSat – an advanced microsatellite based on a microsil bus. *Proceedings of the 12th Annual AIAA/USU Conference on Small Satellites: Smaller than Small – the next generation*. Utah State University, Logan, Utah, USA.

Barton, C.E. (1997). International geomagnetic reference field: the seventh generation. *Journal of Geomagnetism and Geoelectricity*, 49: 121–146.

Chappell, R.C., Moore, T.E. and Waite, J.H., Jr. (1987). The ionosphere as a fully adequate source of plasma for the Earth's magnetosphere. *Journal of Geophysical Research*, 92: 5896–5910.

Elphic, R.C., Means, J.D., Snare, R.C., Strangeway, R.J., Kepko, L. and Ergun, R.E. (2001). Magnetic field instruments for the FAST auroral snapshot explorer. *Space Science Reviews*. (in press).

Erlandson, R.E., Zanetti, L.J. and Potemra, T.A. (1989). Magnetic fluctuations from 0–26 Hz observed from a polar orbiting satellite. *IEEE Transactions on Plasma Science*.

Essex, E.A., Webb, P.A., Horvath, I., McKinnon, C., Shilo, N. and Tate, B. (1999). Monitoring the ionosphere/plasmasphere with low Earth orbit satellites: the Australian microsatellite FedSat. In: Morris, R.J. and Wilkinson, P.J. (Eds.). *ANARE Reports 146*. Australian Antarctic Division. Pp. 321–328.

Fraser, B.J. (1975). Ionospheric duct propagation and Pc1 source regions. *Journal of Geophysical Research*, 80: 2790–2807.

Fraser, B.J. (1985). Observations of ion cyclotron waves near synchronous orbit and on the ground. *Space Science Reviews*, 42: 357–374.

Fukunishi, H., Takahashi, Y., Kubota, M., Sakanoi, K., Inan, U.S. and Lyons, W.A. (1996). Elves: lightning induced transient luminous events in the lower ionosphere. *Geophysical Research Letters*, 23: 2157–2160.

Graham, E.C. (1999). FedSat: an Australian research microsatellite mission, paper IAA-98-IAAA. 11.1.01, IAF Congress, Melbourne, Australia, 1998.

Gurnett, D.A., Shawhan, S.D., Brice, N.M. and Smith, R.L. (1965). Ion cyclotron whistlers. *Journal of Geophysical Research*, 70: 1665–1688.

Iijima, T. and Potemra, T.A. (1976). The amplitude distribution of field aligned currents at northern high latitudes observed by Triad. *Journal of Geophysical Research*, 81: 2165–2174.

Iyemori, T. and Hayashi, K. (1989). Pc1 micropulsations observed by MAGSAT in the ionospheric F-region. *Journal of Geophysical Research*, 94: 93–100.

McPherron, R.L. (1979). Magnetospheric substorms. *Reviews of Geophysics and Space Physics*, 17: 657–681.

Potemra, T.A., Zanetti, L.J., Bythrow, P.F., Erlandson, R.E., Lundin, R., Block, L.P. and Lindquist, P.A. (1988). Resonant geomagnetic field oscillations and Birkeland currents in the morning sector. *Journal of Geophysical Research*, 93: 2661–2674.

Russell, C.T. and Elphic, R.C. (1978). Initial ISEE magnetometer results: Magnetopause observations. *Space Science Reviews*, 22: 681–715.

Russell, C.T., Snare, R.C., Means, J.D., Pierce, D., Dearborn, D., Larson, M., Barr, G. and Le, G. (1995). The GGS/POLAR magnetic fields investigation. *Space Science Reviews*, 71: 563–582.

Samson, J.C. (1983). The spectral matrix, eigenvalue, and principal components in the analysis of multi-channel geophysical data. *Annales Geophysica*, 1. p. 115.

Sentman, D., Wescott, E.M., Osborne, D.L., Hampton, L.H. and Heavner, M.J. (1995). Preliminary results from the sprites94 aircraft campaign. 1. red sprites. *Geophysical Research Letters*, 22: 1205–1208.

Trakhtengertz, V.Y. and Feldstein, A.Y. (1987). Excitation of small-scale electromagnetic disturbances in the ionospheric Alfvén resonator. *Geomagnetism and Aeronomy*, 27: 271–273.

Weimer, D. R., Geortz, C.K., Gurnett, D.A., Maynard, N.C. and Burch, J.L. (1985). Auroral zone electric fields from DE-1 and –2 at magnetic conjunctions. *Journal of Geophysical Research*, 90: 7479–7494.

Woodman, R.F. and LaHoz, C. (1976). Radar observations of F region irregularities. *Journal of Geophysical Research*, 81: 5447–5466.

30. DAVIS AND CASEY DIGITAL MAGNETIC FIELD DATA – AN AUSTRALIAN ANTARCTIC DIVISION REVIEW

P.I. Roberts, G.B. Burns and R.J. Morris

Atmospheric and Space Physics
Australian Antarctic Division
Kingston Tasmania 7050 Australia
(email: gary.burns@aad.gov.au)

Abstract

Digital magnetometer data have been collected from Davis (68.58°S, 78.0°E) since 16 February 1988 and from Casey (66.28°S, 110.5°E) since 17 January 1989. The processing of the variometer data from Casey has been transferred to the Australian Geological Survey Organisation (AGSO) from 1 January 1999. This paper reports on the processing of the Casey data to the end of 1998, the processing of the Davis data to the end of 1999, the processing software and models used, the known, suspected and possible influences on the station baselines, and the reports that exist covering the period of Australian Antarctic Division responsibility for the Davis and Casey variometer data reduction. Emphasis is given to the baseline determinations for the 1997 to 1999 interval at Casey and Davis stations and the insights that these give, in association with previous results. Intervals when unprocessed Macquarie Island (54.50°S, 159.0°E: 7 September 1987 to 4 January 1994) and Mawson (67.60°S, 62.9°E: 8 January 1991 to 19 January 1994) were collected by arrangement with AGSO, are also noted.

30.1 A brief history of digital variometer data collection by the Australian Antarctic Division

Howard Burton developed the initial data collection software LOGIT in FORTRAN and RT-11 MACRO languages that ran on PDP 11 platforms and stored data in PDP binary format. Lloyd Symons developed the ADAS (Analogue Data Acquisition System) software, in the C language, that runs under MS-DOS on PC platforms and stores data in a proprietary binary format. The ADAS system was officially instigated at all stations on 1 January 1995, and its operation is ongoing (Symons, 1997). All raw variometer data collected in PDP binary format have been converted to ADAS format. Only ADAS-formatted raw variometer data records now exist in the archives. ADAS files can be examined and manipulated on a PC using the program, VIEW. The ADAS and VIEW software is continually evolving. To this date all upgrades have been undertaken by Lloyd Symons. Information on the latest versions can be found at the URL:

http://intranet.antdiv.gov.au/asp/ASP_DAQ/ASP_DAQ.htm.

Variometer data were initially collected at a 10 second sample rate. They have been collected from Davis (68.58°S, 78.0°E) since 16 February 1988 and from Casey (66.28°S, 110.5°E) since 17 January 1989. Raw variometer data, collected at ten second resolution but now converted and available only in ADAS format, were also collected at Macquarie Island (54.50°S, 159.0°E) from 7 September 1987 to 4 January 1994 and from Mawson (67.60°S, 62.9°E) from 8 January 1991 to 19 January 1994 by arrangement with the Bureau of Mineral Resources (now Australian Geological Survey Organisation). AGSO (formerly BMR) has always retained the responsibility for processing variometer data from the Macquarie Island and Mawson sites, but these ten-second resolution data are the highest sample rate digital data available from these sites for some portion of the data collection intervals.

The resolution of raw variometer data was increased to one second at Casey from 11 September 1997 and at Davis from 14 September 1997. The data collection rate was increased to be compatible with international practise.

The data amplitude resolution was increased from 12-bit to 16-bit at Casey from 16 December 1997 and at Davis from 30 January 1998. One extra bit of resolution was used to increase the range from the order of 1000 nT to the order of 2000 nT, eliminating those times when large variation events saturated the resolution of the data collection system. Three extra bits of resolution increased the sensitivity from ~ 0.5 nT per digital unit to ~ 0.06 nT per digital unit.

30.2 Instrumentation

At Casey and Davis, modified EDA FM100C fluxgate magnetometers (variometers), are used to record variations in the Earth's magnetic field in the X (geographic North), Y (geographic East) and Z (vertically downward) directions. The temperature, T, variations inside the coil sensor housing are measured by means of a Doric AD2060 sensor. The analogue outputs (X, Y, Z and T) are amplified and filtered, using a low pass filter with a roll-off from around 15 Hz (Dymond, 1990). These signals are fed into an analogue-to-digital converter card inside a logging PC. Time data is logged by means of external timing signals derived from the Australian Antarctic Division designed Time Decoder Card which is synchronised with IRIG-B serial time code from the GPS based chronometers. Data archiving is achieved by means of regular backups to station based MicroVAXES via a local area network, and to the Australian Antarctic Division's VAX computer via the ANARESAT satellite system. Early in 2000, station MicroVAXES were replaced with Linux based PC's to perform the task of on-station data archiving and automatic data transfers by Kym Newbery and Damian Murphy. Figure 1 shows the interconnection of the logging and timing systems.

The variometers are calibrated by means of regular magnetic absolute measurements using a Declinometer/Inclinometer Magnetometer (DIM) for measuring D and I, a Proton Precession Magnetometer (PPM)

for measuring F, and a Quartz Horizontal Magnetometer (QHM) for measuring H and D. The DIM and PPM are the primary absolute instruments and the QHM is used as a secondary instrument should the DIM fail. Due to failure of the Davis DIM, magnetic absolutes were performed using the QHM from 29 July 1999 until a replacement DIM was sent to Davis in January 2000. The results of the QHM observations at Davis are comparable to that of the DIM, hence no distinction has been made between observations made using the two instruments.

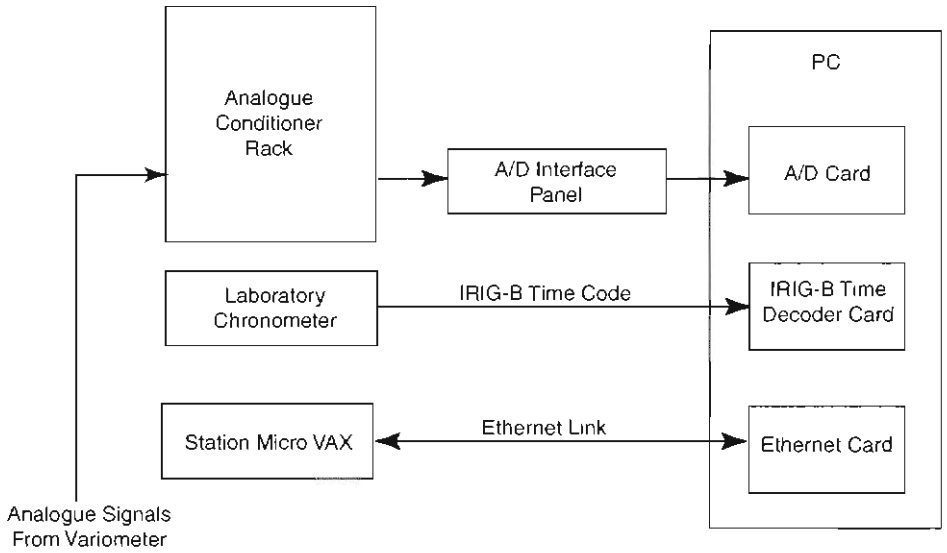


Figure 1. *ADAS instrumentation block diagram.*

30.3 Data format

30.3.1 Raw variometer data

Data files are stored as 24 hour binary files beginning at 00:00:00 UT and ending at 23:59:59 UT. Figure 2 shows an example of a Davis raw ADAS data file for 5 December 1999. For the first minute of each file (i.e. 00:00:01 UT to 00:01:00 UT) and the minute following UT noon (i.e. 12:00:01 UT to 12:01:00 UT), the variometer output of each channel is shorted to provide a baseline value (zero volts out of the variometer). From 16:30:01 UT to 16:31:00 UT a calibration current is passed through a coil, co-located with the variometer sensors, to allow a calibrated deflection on each channel to be logged. Prior to 1995, under the LOGIT system, the zero volt baseline value was provided for one minute on every hour.

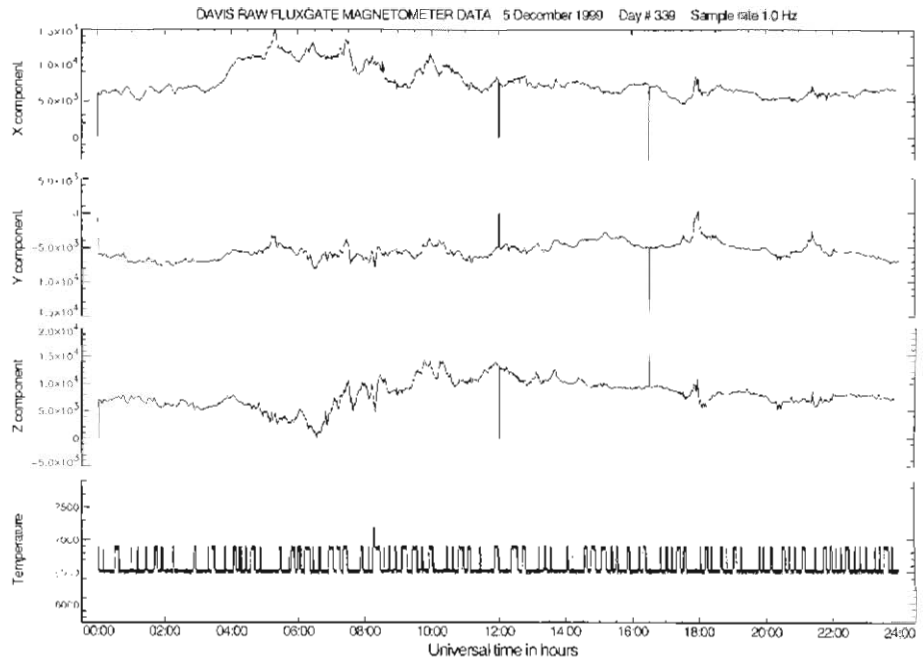


Figure 2. *Raw data file, Davis 5 December 1999.*

30.3.2 Variometer data processing

Barry Giles of Spurion Technology developed a FORTRAN based variometer data analysis program called PROCESS. This initially converted raw data files collected by LOGIT into coded ASCII files of processed data, with a significant volume of added information appended to the file (see Burns and Klekociuk, 1994). An additional FORTRAN program called WDC converted processed files from ten seconds resolution values to one minute averages, and generated a file intended to be in standard format suitable for forwarding to World Data Centre-A (Boulder, Colorado, USA). Stephen Newbery refined the PROCESS program and corrected the WDC code. These programs are described in detail by a range of Australian Antarctic Division internal reports listed in Appendix F. The development of these programs occurred during 1992 and 1993.

The program WDC also has the facility for generating hourly average values in World Data Centre format, but the output of this option has never been fully tested. It was decided, through correspondence with WDC-A, that the Casey and Davis sites did not have their baselines sufficiently defined (absolute measurements were not conducted often enough) for the sites to be of 'observatory-quality'. WDC-A only collects hourly average variometer data from observatory-quality sites.

Prior to completion of the back-processing of all digital variometer data into WDC-A one-minute resolution format, Volodya Papitashvili, who was familiar with the World Data Centre magnetic field codes, noted that the output formulation of the minute-averages by the WDC program was not fully consistent. He developed a 16-bit PC based program (MVFIX_AU) and a procedure (see Appendix E) for correcting the mis-codings.

Campaigns were occasionally conducted with variometer data collected at one second resolution prior to the dates when this resolution became the standard operational procedure. Malcolm Lambert developed a version of PROCESS (called PROCESS_SEC) which generates one second resolution processed data from the one second resolution LOGIT data. This program contains an option for generating ten second resolution processed files for input to WDC.

The instigation of PC-based variometer data collection via ADAS in 1995 led to the development by Didier Monselesan of IDL based programs for processing the new raw variometer data files. These IDL routines have been developed to produce processed files suitable for input to the WDC programs from ten or one second resolution ADAS data, 16-bit or 12-bit resolution data, with an option for generating processed data. The processed files produced in this manner do not contain the full range of additional information provided by the PROCESS.FOR routines (see Appendix B for the IDL output format, and compare with the PROCESS outputs presented in Burns and Klekociuk, 1994). These IDL routines were developed as a temporary means to avoid the difficulty the ADAS data posed to the production of the World Data Centre formatted one-minute averages.

It was initially planned that Perry Roberts would develop a variometer data processing package that eliminated the contortions that had developed in the procedure (see Appendix E). Subsequent to his appointment, the Geomagnetism Section of AGSO received additional funding which included provision for upgrading the Casey site to full observatory status. Control of the Casey variometer data was transferred to AGSO from 1 January 1999. It is hoped that control of Davis variometer data can also be negotiated. Perry Robert's project evolved to processing the 1997 to 1999 variometer data that was the responsibility of the Australian Antarctic Division, and upgrading and unifying the absolutes reduction procedures at Davis and Casey.

When the raw data is processed a variometer model is established (see Appendix A). This model is used to process the binary variometer files to ASCII format, an example of which is shown in Appendix B. The 0000 UT, 1200 UT and the 1630 UT marks, along with any undesirable data spikes, are removed during the processing and replaced with 999999999999999 ASCII characters. Any data gaps in the raw data are similarly replaced. Time information of any given data record can be determined from its location in the file.

The first record in the file is the file name and the second record specifies the date at which the ASCII file was created. Record three is offset baseline values for X, Y and Z. The next 8640 records are coded data values in the order X, Y and Z. This represents a data point every ten seconds irrespective as to whether the raw binary files were logged as one-second or ten-second data. The remaining records are the calibration equations used as defined in Section 30.5. Prior to 1995 this section of the file contained additional information relevant to that days data (see Burns and Klekociuk, 1994).

The ASCII files are further processed to produce one-minute data files in a World Data Centre (WDC) format. An example of a fully processed single days data is shown in Figure 3.

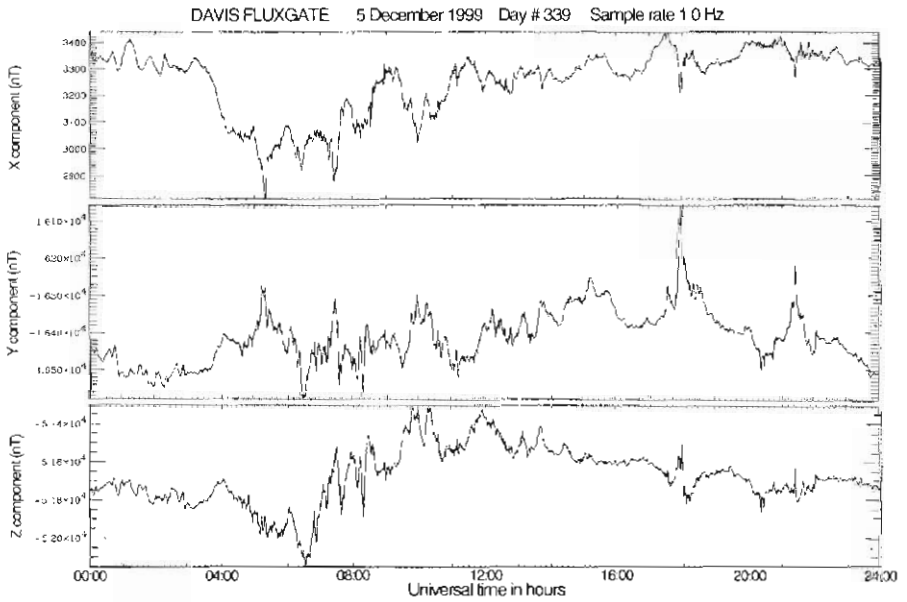


Figure 3. *Processed data file, Davis 5 December 1999.*

30.4 Casey and Davis calibrated data, 1992–99

30.4.1 Discussion

The data plots in this section show the variation in the calibrated values of X, Y, Z, H, F, D and I for the period 1997 to 1999 inclusive. Figures 4 to 10 show the plots for Casey calibrated data and Figures 12 to 18 show the plots for Davis calibrated data. Plots of the temperature, T, variation of the variometer sensor coils, for the period 1997 to 1999, are shown in Figures 11 and 19 for Casey and Davis respectively.

The calibrated values are determined by the reduction and comparison of magnetic absolute observations with variometer data. The reduction process produces baseline residuals which are corrections to the variometer's predicted values of X, Y and Z to make them agree with the values of D, I and F as measured with the absolute instruments. A baseline model is determined from the averages of the calibrated values calculated over selected periods for a given year. The data in Appendix A are the baseline models that have been adopted at Casey since 1989 and at Davis since 1988. For data 1997 to 1998, at Casey, the range of baseline values amount to 35 nT, 45 nT and 14 nT for the X, Y and Z channels respectively. For Davis, 1997 to 1999, the range of baselines amount to 59 nT, 39 nT and 36 nT for the X, Y and Z channels respectively. The baseline data for Casey 1999 were not included in the analysis as the responsibility for this has been transferred to AGSO.

The values in Table 1 are the standard deviations of the pre-averaged data shown in the plots. In the analysis, 302 and 144 data points were used for Casey and Davis respectively. Davis X and Z channels have consistently higher scatter in baseline values compared with Casey data. The differences in scatter of the Y channel is less marked, but is higher for Casey compared with Davis. The use of standard deviations alone does not give a true indication of the amount of scatter in the data set.

Table1. *Casey and Davis standard deviation data, 1997 to 1999.*

Casey Standard Deviation Data

Year (nT)	X (nT)	Y (nT)	Z (nT)	H (nT)	F (degrees)	D (degrees)	I	T (°C)
1997	11.9	10.7	3.1	10.7	3.4	0.07	0.01	0.21
1998	7.4	17.0	4.7	17.2	6.2	0.04	0.01	0.20
1999	10.3	12.2	7.0	12.3	7.8	0.06	0.01	0.23

Davis Standard Deviation Data

Year (nT)	X (nT)	Y (nT)	Z (nT)	H (nT)	F (degrees)	D (degrees)	I	T (°C)
1997	20.7	8.9	11.2	12.5	10.0	0.07	0.01	0.18
1998	22.6	8.9	9.3	12.8	6.6	0.07	0.02	0.20
1999	18.5	10.6	8.2	14.3	8.4	0.05	0.01	0.22

For Davis, an examination of the data plots of X, Y and Z channels show a seasonal variation, particularly in the horizontal plane (see Figure 15). This variation was first observed in analysis of magnetic data from 1990. "We now suspect that the freezing and thawing of the 'gravel' on which the concrete pad for the magnetometers sensors is located is causing a seasonal tilt in the sensors" (Burns and Klekociuk, 1994). The seasonal variation for the 1997 to 1999 Davis data has a similar profile to the 1990 data discussed by Burns and Klekociuk (1994). That is, during the summer periods there is an apparent gradual change in baselines levels to a winter level for which the baseline remains less variable for a much longer period of time. As pointed out by Stewart Dennis, AGSO, when reviewing this paper, if the seasonal baseline drift at Davis is due to a 'tilt' process, then the F baseline should be maintained, while the Z and H baselines should vary in a complementary manner. The sense and approximate magnitudes of the Z and H baseline variations at Davis are not inconsistent with a seasonal tilt. The Davis F baselines show an ~ 30 nT discontinuity in early 1998 but otherwise a long-term rather than a seasonal variation. The noise in the measurements precludes a definitive conclusion.

The Y component of Casey data for 1997 (see Figure 5) shows a similar seasonal variation. From February 1999, the numbers of absolute observations performed at Casey were increased to around two per week providing a clearer indication of the baseline variation. While there is some hint of the seasonal variation in the 1997 and 1998 data they are not conclusive. The Casey variometer site was relocated, in 1992, from the old 'tunnel site' to 150 m west of the new science building. Prior to the processing and analysis of the 1999 data, a seasonal variation of magnetic data at Casey was not anticipated. At Casey, the H and F baselines in 1999 show a similar shaped seasonal variation. This is inconsistent with the principal cause being a seasonal 'tilt'. Stewart Dennis suggests that seasonal temperature influences cannot be excluded by accurate maintenance of temperature at the sensor heads or manufacturer's claims. He reports that Canadian magneticians tracked a diurnal variation to a section of cable being in and out of the sun for parts of the day.

It has been assumed that the temperature of the sensor coils remains stable and hence variations in the variometer X, Y, Z field components with temperature can be ignored. The data of Figures 11 and 19 indicate that this is a reasonable assumption. For the period 1997 to 1999 the average sensor temperature for Casey is 20.6°C over the range 19.9°C to 21.4°C (a 1.6°C spread). For Davis the average is 20.1°C over the range 19.8°C to 20.7°C (a 0.8°C spread). The manufacturers of the EDA-100C magnetometers state that temperature sensitivity is 1 nT/°C. Hence temperature variations of the coils would produce a negligible (less than 2 nT) change in the measured field.

If the seasonal variation in the Casey and Davis data is taken into account, the scatter in the data sets will be significantly less than

indicated by the standard deviation figures in Table 1. To illustrate this the Casey data was analysed using 110 data points between 12 April 1999 (Day 102) and 26 October 1999 (Day 299). This period is roughly the winter interval in the seasonal variation. The standard deviations for this period are 5.7 nT, 6.1 nT and 3.6 nT for the X, Y and Z baselines respectively. As a comparison, the 1999 magnetic data for Macquarie Island was processed using a fixed baseline reference for the entire year's data. The standard deviations for the Macquarie Island X, Y and Z channels are 1.9 nT, 1.7 nT and 0.9 nT respectively. While the Casey 1999 winter data compares more favourably with the Macquarie Island data, there is still significantly more scatter in the Casey winter data.

There are physical differences in the geomagnetic systems of Casey, Davis and Macquarie Island that may account for variations across data sets. All three stations use the same type of DIM as the primary absolute instrument, with the exception that at Davis a QHM was used from 29 July 1999 until a replacement DIM was sent to Davis in January 2000. The absolute piers on which these instruments sit are not identical. The materials used and construction methods vary, as do the number of piers. At Davis the same pier is used for PPM and DIM or QHM observations. At Macquarie Island PPM observations are made on a different pier to DIM and QHM observations. Prior to 1993, Casey also used separate piers for PPM, DIM or QHM observations. There is an approximate 423 nT pier difference in F between the two piers (McLoughlin, 1990). While only one pier is now used, the gradient indicates a region of significant magnetic variability in the vicinity of the Casey absolutes site. There is also a vertical gradient in the absolutes pier which had to be accounted for when the DIM replaced by the QHM as the primary absolute instrument. The DIM's effective magnetic centre, defined by the centre of the DIM telescope, is 8 cm higher than the QHM's effective magnetic centre (Symons and Dubovinsky, 1994). In order for the DIM/PPM measurements to be directly compared with the QHM/PPM measurements an effective height correction is added to all DIM/PPM results. The effective correction, determined from measurements over the 1992–93 and 1994–95 summers (Symons and Dubovinsky, 1994), amount to a +15.1′ in D, +0.2′ in I and +45 nT in F. This translates to +42 nT, -11.5 nT and -44.0 nT in X, Y and Z respectively. These corrections have been applied to all magnetic data presented in this report.

Magnetic absolute observations have an associated observation error. For the DIM observations this depends on the levelling accuracy of the instrument, imperfections in the theodolite and errors in horizontal and vertical circle readings. Errors due to instrument imperfections are minimised by calibrations against standard instruments. The possible error introduced due to variations in the accuracy to which the DIM is levelled is difficult to quantify but it is thought to be small relative to the overall uncertainty in the data. The errors in horizontal and vertical circle readings, however can be quantified. The DIM's at Casey and Davis have a measurement accuracy of 0.1 minute. The data in Tables 2 and 3 of Appendix D show the effect of an error in declination and inclination

on the baselines of X, Y and Z for Casey, Davis, Mawson and Macquarie Island. For an observation error in declination of 0.1 minute the worst case variation in the baselines of X, Y and Z at Casey or Davis is 0.5 nT. For inclination this figure is 1.9 nT. For an observation error 10 times the accuracy of the DIM (i.e. 1 minute) the worst case error is 4.9 nT due to a declination error and 18.6 nT due to an inclination error. The magnitude of the seasonal variation for Casey 1999 baseline data is roughly 30 nT, 40 nT and 20 nT for the X, Y and Z components respectively. For Davis these figures are 40 nT, 35 nT and 25 nT. Hence observation errors cannot account for the seasonal variation observed. Differences between stations of the X, Y and Z baselines to the sensitivity of the D, I and F absolute readings also cannot account for the larger standard deviations in the Casey and Davis absolute measurements.

The variometers in use at the stations also differ. Casey and Davis use EDA Fluxgate Magnetometers whereas Macquarie Island and Mawson use Ring Core Fluxgate Magnetometers. At each site the electronics for the variometers are kept under laboratory conditions in which temperature variations are minimised so the effect of temperature on the variometer electronics is not considered a source of the seasonal variation. However some differences in the general spread and accuracy of the variometer data may be attributable to the different styles of variometers. The sensor coils at the stations are installed using different methods. At Casey and Davis the coils are mounted on non-magnetic concrete pads. At Davis the pad rests on the gravelly ground and at Casey the pad is attached to a large boulder. Additionally the Casey site is near the station quarry which over the years has been subject to significant human activity, including the use of explosives. This may have contributed to some unsettling of the variometer coil foundation. While at both Casey and Davis the coils are in a heated housing, the concrete pads on which they rest are exposed to the environment. Contrast this with the Macquarie Island and Mawson observatories which have the variometer coils resting on concrete piers inside heated buildings. With the piers at Casey and Davis having a considerable exposure to the extremes of temperatures, and in particular the potential for a freezing and thawing of the surrounding ground, a mechanism for the observed seasonal variation in the magnetic data does exist. It is also possible, that even during the winter stage of the seasonal variation, a pier instability exists. This may contribute to the general scatter of the data.

In a 1992 experiment at Davis, a high resolution quartz magnetometer was compared with the Davis EDA fluxgate magnetometer (see Paptashvili *et al.*, 1996). A major concern of this paper was the correct orientation of the magnetic instruments against the geographic or geomagnetic axes, and to a local vertical line. A conclusion was that large values of the geomagnetic Z component introduce significant errors in the horizontal field components if the local vertical line is not properly defined. The figures quoted, namely, about 90 nT for each 0.1° at Davis and 110 nT at Casey are in line with the results presented in

Table 3 of Appendix D. The results presented in the paper indicate the magnitude of the local variations in the mutual vertical alignment of the two instruments is in the range -6.5° to +5.0°. However the magnitude of these angles appears unrealistically large when considering the magnitude of the magnetic variations observed (less than 100 nT) in the data.

The paper also recommends the installation of tilt meters, with an accuracy of 0.01°, on the concrete piers of the Casey and Davis variometers. That recommendation is supported by this paper as it may provide a correlation with the seasonal variation observed in the magnetic data. It should be appreciated that the magnitude, of the seasonal variation (20 nT–40 nT), may be produced by tilts as small as 0.02°. The upgrade or replacement of existing variometer piers at Casey and Davis could also be considered, to bring them in line with the standards achieved at the Macquarie Island and Mawson sites.

It has been assumed that the ASP scale values, used in the calibration equations, are accurate. The stability of the scale values can be tested by a comparison of 1630 UT deflections, after accounting for the calibration current through the coils. The magnitude of the scale values have only been determined once at each site, so it is recommended that the scale values be re-determined. Ideally this would be done by an intensive campaign of absolute observations over a short period of time, preferably in an active magnetic field.

30.4.2 Casey plots

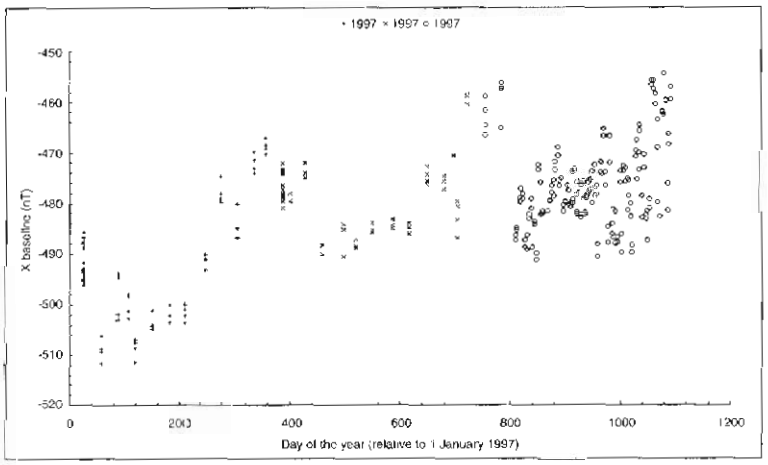


Figure 4. *Casey variometer X component baseline.*

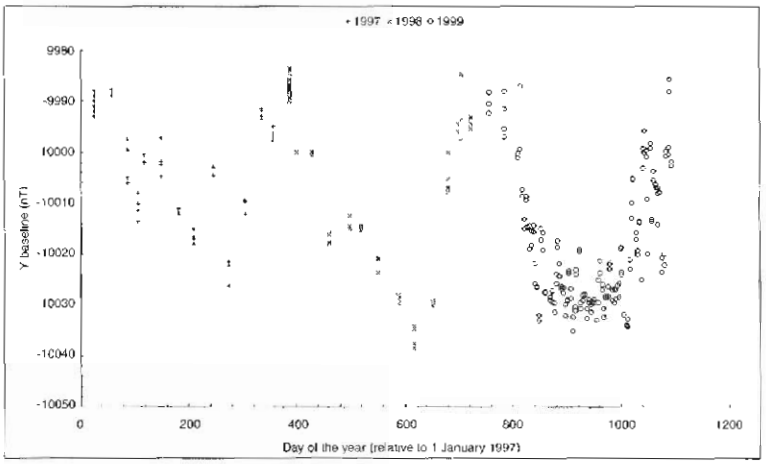


Figure 5. *Casey variometer Y component baseline.*

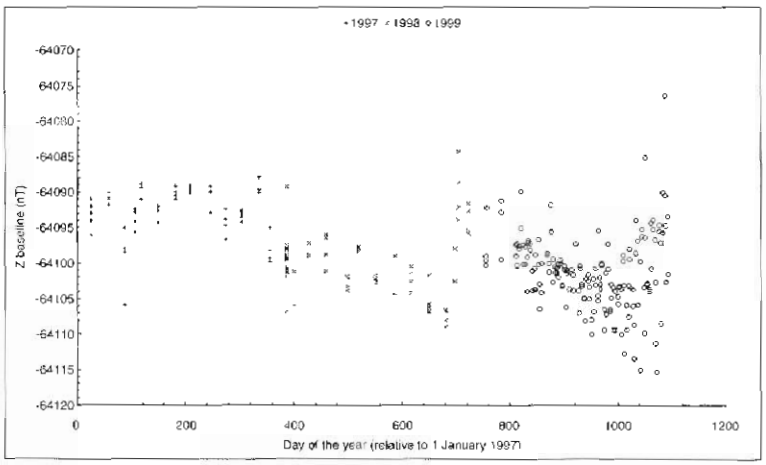


Figure 6. *Casey variometer Z component baseline.*

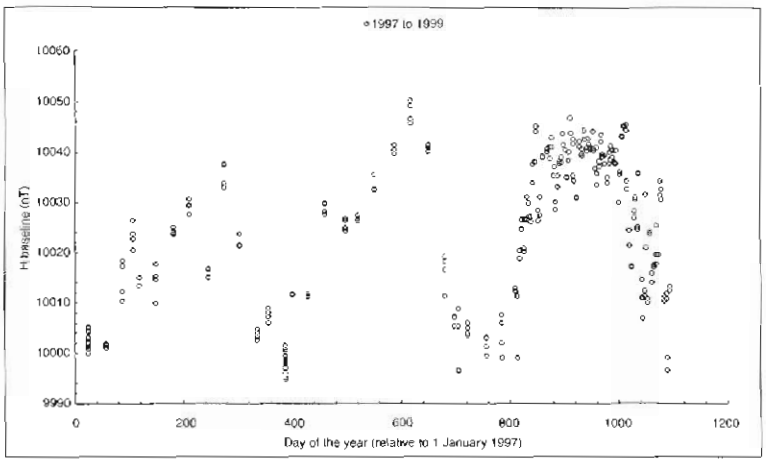


Figure 7. *Casey variometer horizontal component baseline.*

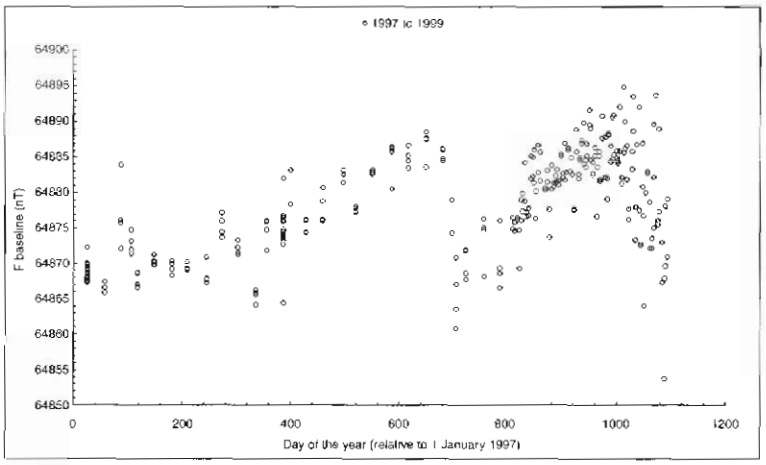


Figure 8. *Casey variometer total field baseline.*

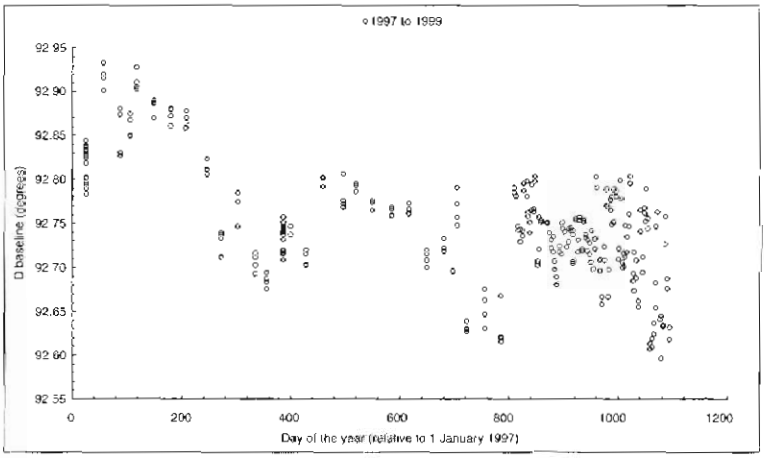


Figure 9. *Casey variometer declination baseline.*

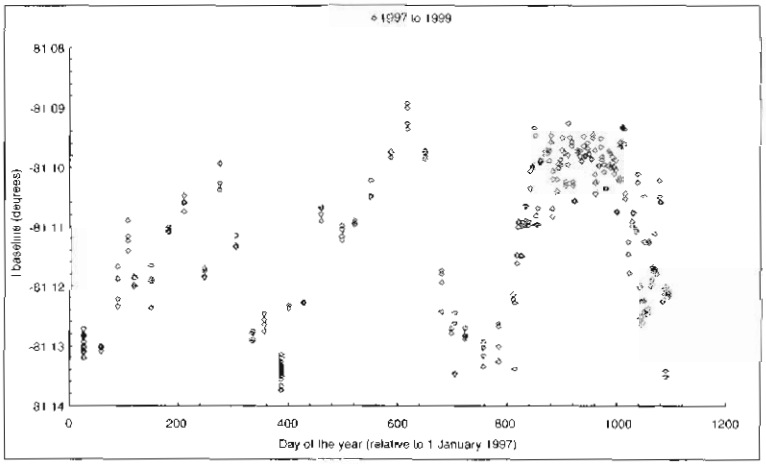


Figure 10. *Casey variometer inclination baseline.*

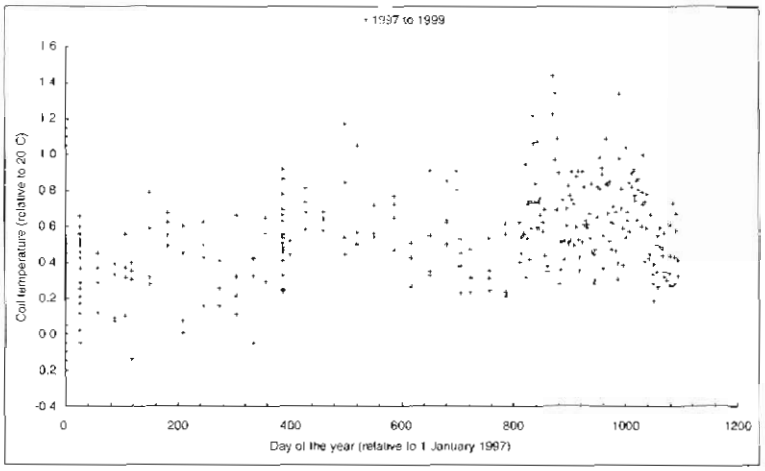


Figure 11. *Casey variometer coil temperature.*

30.4.3 Davis plots

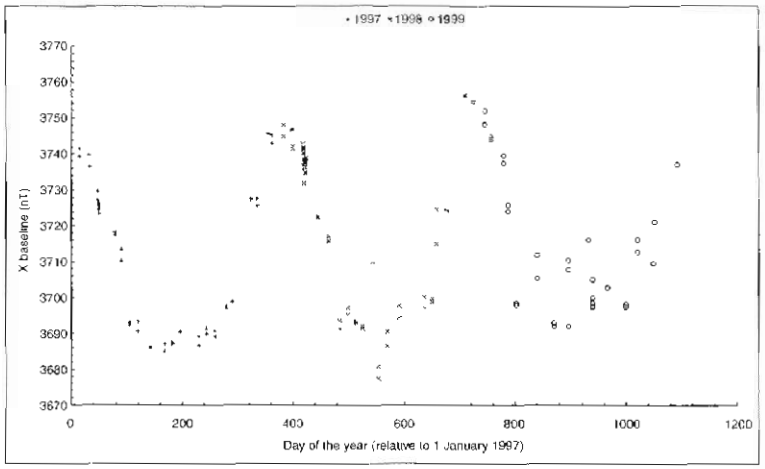


Figure 12. *Davis variometer X component baseline.*

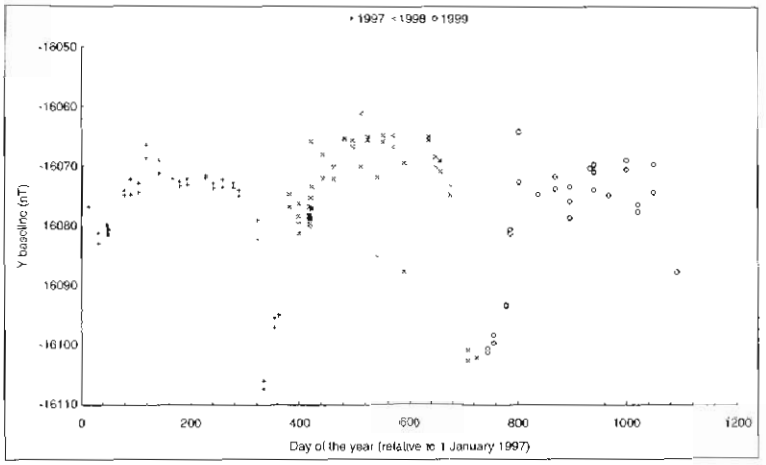


Figure 13. *Davis variometer Y component baseline.*

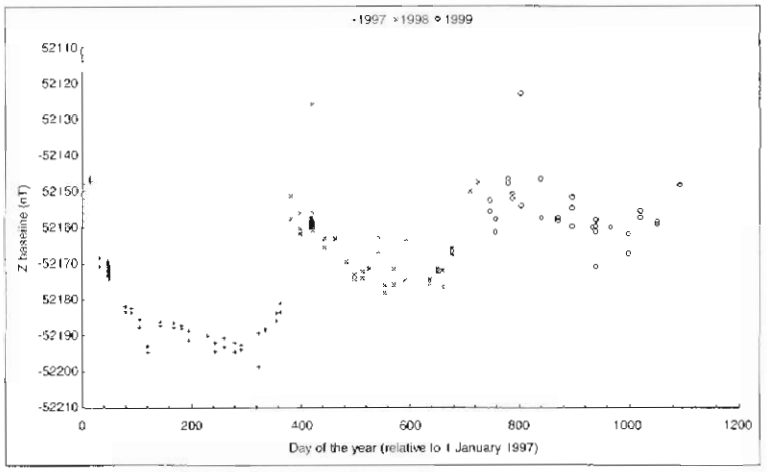


Figure 14. *Davis variometer Z component baseline.*

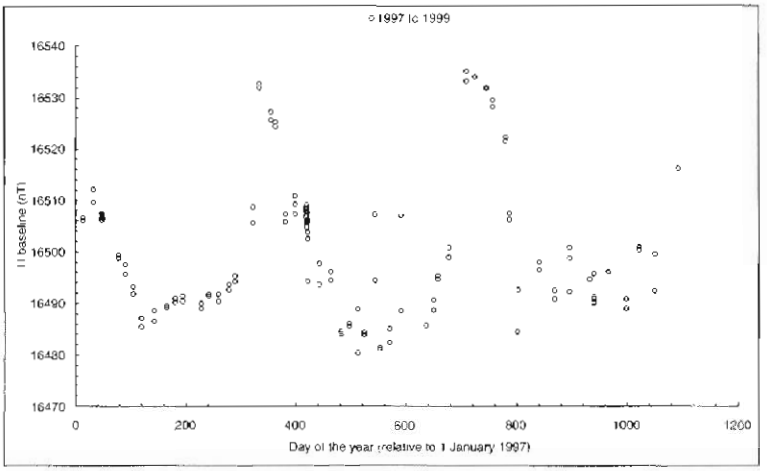


Figure 15. *Davis variometer horizontal component baseline.*

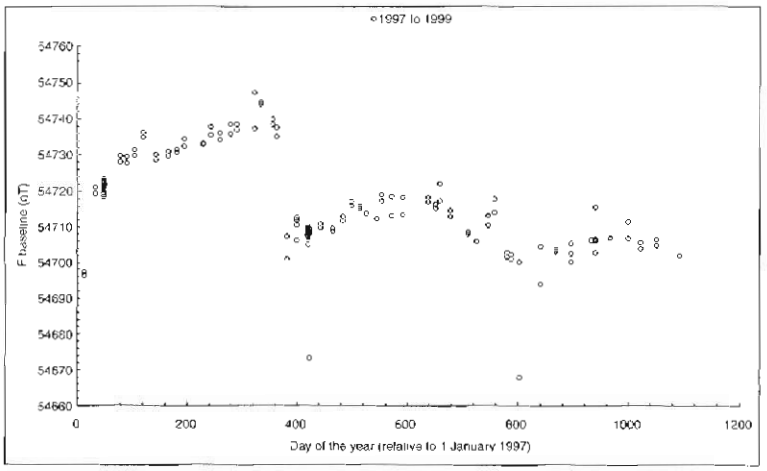


Figure 16. *Davis variometer total field baseline.*

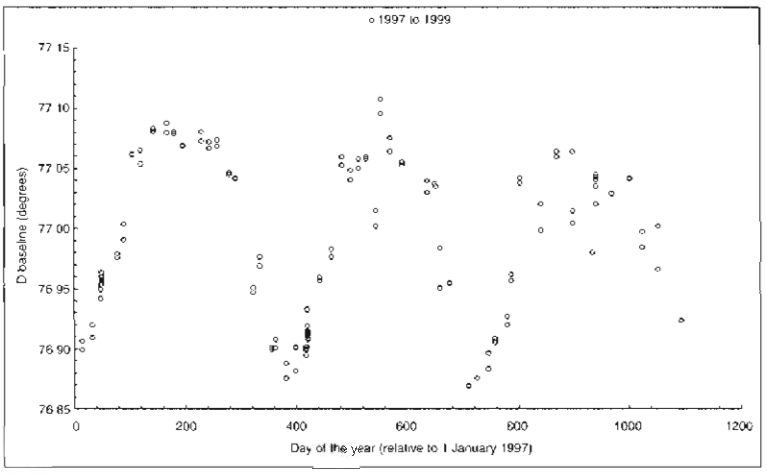


Figure 17. *Davis variometer declination baseline.*

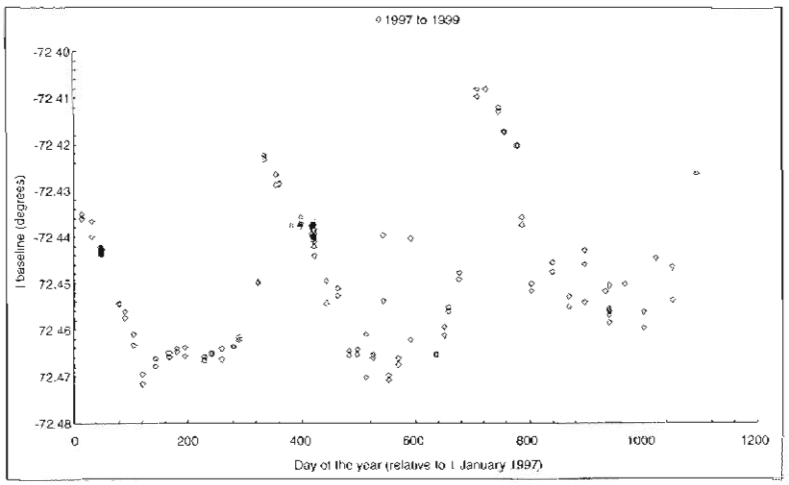


Figure 18. *Davis variometer inclination baseline.*

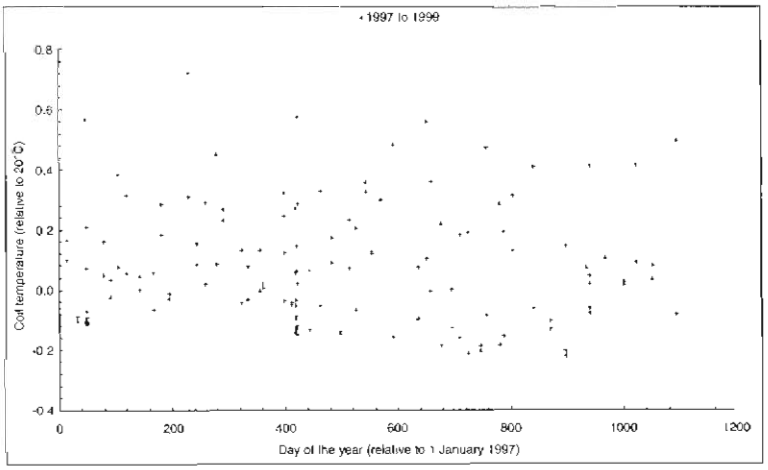


Figure 19. *Davis variometer coil temperature.*

30.5 Calibration system equations

To calibrate the variometer system the output voltage of each channel, which are related to the digital counts logged at the computer interface, is converted into a nT value. The calibration process is represented by the equations below.

30.5.1 For the X, Y, and Z channels

The magnetic field intensity, for the X, Y and Z channels is given by:

$$B = Sasp*V + BLasp \quad nT \qquad (1)$$

where,

Sasp = ASP scale values or sensitivities of the variometer channels in nT/volt.

V = voltage outputs from the variometer sensors in volts.

BLasp = ASP baseline values for the variometer in nT.

The voltage outputs from the variometer sensors can be expressed as:

$$V = (D + Offset)/(A*Gain) \quad volts \qquad (2)$$

where,

D = logged digital values representing B in digital counts.

Offset = logged digital values corresponding to zero volts out of the variometer in digital counts.

A = digital counts per voltage range in digital counts/volt.

Gain = amplification factor for the analogue conditioner channels.

Combining Equations (1) and (2) we obtain:

$$B = [Sasp/(A*Gain)]*D + Sasp*[Offset/(A*Gain)] + BLasp \quad nT \qquad (3)$$

30.5.2 For the temperature channel

The equations for the temperature channel are similar to those for the X, Y and Z channels. Hence only the final equation is shown here.

The temperature, as measured by the Doric Thermometer, in the variometer coil housing is given by:

$$T = [Sasp/(A*Gain)]*D + Sasp*[Offset/(A*Gain)] + TLasp \quad T°C \qquad (4)$$

where,

Sasp = ASP scale values or sensitivities of the temperature channel in nT/volt.

A = digital counts per voltage range in digital counts/volt.

Gain = amplification factor for the analogue conditioner channel.

D = logged digital values representing T in digital counts.

Offset = logged digital values corresponding to zero volts out of the Doric Thermometer in digital counts.

TLasp = ASP baseline value for the Doric Thermometer in °C.

30.5.3 Relationship between asp and AGSO system equations

Baseline values are determined by regular magnetic absolute observations. The use of AGSO software to reduce these observations to field coordinates requires a correction factor to be added to the results in order to convert them to ASP baselines.

The equivalent AGSO system equation for Equation (1) is given by:

$$B = Sagso*D + BLagso \quad nT \tag{5}$$

where,

Sagso = AGSO scale values or sensitivities of the variometer channels in nT/digital count.

D = logged digital values representing B in digital counts.

BLagso = AGSO baseline values for the variometer in nT.

If Equations (3) and (5) are combined we obtain:

$$Sagso = [Sasp/(A*Gain)] \quad nT/count \tag{6}$$

and

$$BLagso = Sasp*[Offset/(A*Gain)] + BLasp \quad nT \tag{7}$$

30.6 Processing software

30.6.1 The current system

Processing magnetic absolute data requires the use of software routines and input files to reduce the absolute data to field components. This reduced data is then compared with the variometer data as part of the calibration process. A list of software used for reduction and comparison is shown below.

Casey only:
CASABS.BAT (main batch file)
MAGINFO.CSY or MAGINFO.CAS (instrument and site corrections file)

Davis only:
DAVABS.BAT (main batch file)
MAGINFO.DAV (instrument and site corrections file)

Common files:
INTRM.iii (instrument component file, where iii is an id number representing a particular DIM, QHM or PPM)
VIEW.EXE (ADAS view program)
REDUCE.EXE (AGSO program)
COMPARE.EXE (AGSO program)
MACQOBS.EXE (AGSO program)
MACQCSV.EXE (AGSO program)

Each set of absolute observations is typed up into a format acceptable to the processing software. This file is called the raw absolutes file and always has the file extension .OB. The batch file CASABS.BAT (for Casey) or DAVABS.BAT (for Davis) is the main executable and all other files are called from within this batch file. Initially the binary variometer data file is converted to an ASCII format using VIEW.EXE. When the main batch file is executed the absolute observations (in the .OB file) are merged with the ASCII variometer data at the corresponding absolute observation times. The merged data file (file extension .OBS) is then reduced to field components F, D and I for PPM/DIM observations and to F, H and D for PPM/QHM observations. Each absolute observation is adjusted to a field value, calculated by means of averaging the variations in the magnetic field as measured by the variometer, over the period of the absolute observation. The reduced data are then compared with the variometer data to provide baseline residual information. Usually once per year, baseline values for X, Y and Z are determined by performing a number of absolute observations (around 10 PPM/DIM sets) over a short period of time. This produces a set of baselines to act as a reference for the variometer. The residual information obtained each month is added to the annual baseline values to allow any steps or drifts in the variometer's baselines to be monitored.

An important input file for each station is the instrument and corrections files MAGINFO.CSY (formerly MAGINFO.CAS) and MAGINFO.DAV. These files contain instrument, site and station information used by the software. Part of the content of these files is shown in Appendix C. These data are calculated as AGSO values as opposed to the ASP values. This is necessary as the AGSO software expects data based on the AGSO form of the magnetic model.

The basic difference in the two models is that the ASP model takes the baselines as 'zero volts out of the variometer' and the AGSO model uses 'zero digital units' as the baselines. An example (using Casey 1997, Day 354) of each step in the reduction and comparison process is shown below.

Merge the raw absolutes file with the ASCII version of the days ADAS data file.
Input file #1: CAS97354.TXT (ASCII version of the ADAS data file)
Input file #2: C97354D.OB (raw absolutes file)
Executable: MACCQBS.EXE
Instrument File: nil
Output File: C97354D.OBS (merged data and absolutes file)

Reduce the absolute observations to field components:
Input file: C97354D.OBS (merged data and absolutes file)
Executable: REDUCE.EXE
Instrument File: MAGINFO.CAS
Output File: C97354D.RED (reduced absolutes/data file)

Compare the results of the absolute observations and the variometer measurements:
Input file: C97354D.RED (reduced absolutes/data file)
Executable: COMPARE.EXE
Instrument Files: MAGINFO.CAS, INSTRM.593
Output File: C97354D.BLV (baseline file)

Optional: create a comma separated value (CSV) text file:
Input file: C97354D.BLV (baseline file)
Executable: MACQCSV.EXE
Instrument File: MAGINFO.CAS
Output File: C97354D.CSV (comma separated value file)

30.6.2 The future direction

Prior to 1999, the method adopted for the reduction of magnetic absolutes data at Casey and Davis was the responsibility of the station observers. Subsequent reprocessing and analysis of absolute data for 1996 through 1999 has shown that a number of errors had crept into the system meaning that in the past some magnetic absolute data has been processed with incorrect instrument and baseline information.

The data reduction system used is a combination of software developed by AGSO and individual observers. The overall magnetic absolute system is quite complex to the extent that there are a significant number of stages, utilising various software routines and input files, that are required to correctly reduce the absolute data to field components. An added difficulty was that during 1999 AGSO took over responsibility for magnetic absolutes at Casey as part of the upgrade of Casey to an official observatory. While this has long term benefits the transition from the old system to an AGSO-like system increased the complexity.

During 1999 the magnetic absolute reduction process at Casey and Davis was standardised to reduce the need for observers to write their own processing software and modify system files. This has proved to be relatively successful, and data reduction at the stations has been streamlined and errors minimised. Where possible past errors have been corrected for the 1997 through 2000 absolutes data. While this new system for processing and managing magnetic absolutes is an improvement over past years there is still significant room for further improvement.

Currently critical processing files are:
CASABS.BAT (main Casey batch file)
MAGINFO.CSY or MAGINFO.CAS (instrument and site corrections file)
DAVABS.BAT (main Davis batch file)
MAGINFO.DAV (instrument and site corrections file)
INTRM.iii (instrument component file, where iii is an id number representing a particular DIM, QHM or PPM)

All other files required are AGSO written executables, which have proved to be robust and relatively bug free.

Each time a change occurs, such as a new instrument or an annual baseline calibration, the above files require modification. History has shown that putting the onus on expeditioner staff to make such changes results in inefficiencies and errors in the reduction process. This directly affects data quality. AGSO is the predominant caretaker of the station's magnetic data, hence it seems appropriate that AGSO take over the role of managing the overall absolute systems at all four stations. The need for someone to coordinate updates and modifications to instrument files and supporting software would be better done by AGSO.

There is no doubt that a fully generic approach to the data reduction process would require some effort, but the benefits to data quality and improved efficiency of on-station processing would likely justify these efforts. A software system to further streamline data reductions which could be used as a model for Casey and Davis is suggested below. Ideally the software processing system should also encompass Mawson and Macquarie Island.

It is suggested that the existing batch files CASABS.BAT and DAVABS.BAT be replaced with a single batch file MAGABS.BAT. This 'top level' file would be used to execute all other files and perform actions such as error checking and copying proformas. To simplify the file structure of MAGABS.BAT a second batch file (INFO.BAT) is recommended to allow station, date, year, century, observation type (DIM or QHM) and instrument id information to be identified and accessed by MAGABS.BAT. This would allow information to be changed on the fly without a need to modify MAGABS.BAT. A similar system to this is currently used at Mawson and Macquarie Island. It would be beneficial to write MAGABS.BAT as a windows based program (e.g. written in one of the VISUAL packages) thus providing a more user friendly product. However a DOS based batch file system does allow upgrades without the need for significant programming efforts.

Raw observation data are entered into one of two proformas, namely, SYYdddD.OB or SYYdddQ.OB (where, S is the station code C or D,, YY is the year, ddd is the day number, D or Q identifies the observation as a DIM or QHM and .OB is the field specification for a raw observation file). Currently these files are created by manually copying and renaming master versions of the proformas. It is proposed that MAGABS.BAT automatically create and name these files. Observation data can then be entered as per normal.

INSINFO.CON is the master instruments and corrections file AGSO currently maintains. It contains the instrument id numbers and corrections for all instruments used in the AGSO observatory network (including Casey and Davis). Unfortunately station and pier data is not contained in this file so it could not be used with MAGABS.BAT without modification. With suitable modification it could replace both MAGINFO.CSY and MAGINFO.DAV. This would allow AGSO to issue to the stations an updated version of a single master file should changes occur.

It is recommended a new file SSSINFO.CON (where SSS is the three letter station id) be implemented. This file would be station specific and contain variometer baseline information used by COMPARE.EXE. Currently this information is combined with the instrument and corrections files MAGINFO.CSY and MAGINFO.DAV.

INTRM.iii (where iii is part of the id number for a given DIM or QHM instrument) is a file containing fundamental component, station and instrument information used by REDUCE.EXE. This file is part of the current system.

REDUCE.EXE, COMPARE.EXE, MACQOBS.EXE and MACQCSV.EXE are existing AGSO files that perform the bulk of the processing. These are currently upgraded by AGSO as the need arises.

30.7 Acknowledgements

The data acquisition software used over the years, written by Howard Burton (LOGIT) and Lloyd Symons (ADAS), has proved to be very reliable. The software used to process the data to WDC-A standards was written at various stages by Barry Giles, Stephen Newbery, Malcolm Lambert and Didier Monselesan.

Thanks to Stewart Dennis (AGSO) and Peter Crosthwaite (AGSO) for supplying much of the software used in the reduction and analysis of the absolute data.

The collection of Antarctic data is facilitated by Antarctic Division personnel and ANARE expeditioners. Their assistance is gratefully acknowledged.

Particular recognition goes to the many expeditioners who performed the magnetic absolute observations from 1997 to 1999. At Casey the magnetic observers were Darryn Schneider, Anthony Breed, Simon Pearce, Todd Maddern and Steve Wallace. At Davis the magnetic observers were Roland Leschinski, Mike Manion, Andrew Bish and Lloyd Symons.

Thanks also to Vladimir Papitashvili for his continued interest and assistance in the Casey and Davis magnetometer data set, particularly for checking the variometer data prior to sending it to the World Data Centre on ASP's behalf.

References

Burns, G.B. and Klekociuk, A.R. (1994). Davis and Casey Magnetic Field Data. In: Morris, R.J. (Ed.). *ANARE Research Notes 92*. Pp. 17–41.

Crosthwaite, P. (1994). Personal Computer Data Acquisition (MACQ) and Reduction Software for Geomagnetic Observatories: Reference Manual. *Geomagnetism Note 1994–12*.

Dymond, M. (1990). Analogue Conditioner Rack Technical Manual. *Internal Report.* Australian Antarctic Division.

McLoughlin, R. (1990). Casey Magnetometer Calibration 20/07/90. *Internal Report.* Australian Antarctic Division.

Paptashvili, V.O., Belov, B.A., Burtsev, Yu. A., Pimenov, I.A., Burns, G.B., Morris, R.J., Neudegg, D.A., and Symons, L.P. (1996). Comparison of High Resolution Quartz and Fluxgate Magnetometer Data Recorded at Davis, Antarctica. In: Morris, R.J. (Ed.). *ANARE Research Notes 95.* Australian Antarctic Division. Pp. 81–92.

Symons, L. (1997). Users Guide to ADAS. Version 4.0. *Internal Report.* Australian Antarctic Division.

Symons, L.P. and Dubovinsky, M. (1994). Casey Magnetometer Calibration January 1994. *Internal Report.* Australian Antarctic Division.

APPENDIX A
VARIOMETER MODEL

The following equations represent the baseline model for the Casey EDA variometer for years 1989 to 1998 and the Davis EDA variometer for years 1988 to 1999.

A1 Casey X, Y and Z components

1989 01 17 00 00

Bx [nT] = -185.0 [nT/volt] x Vx [volts] -210 [nT]
By [nT] = -204.0 [nT/volt] x Vy [volts] -9667 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64118 [nT]

1989 03 04 00 00

Bx [nT] = -187.0 [nT/volt] x Vx [volts] -210 [nT]
By [nT] = -204.0 [nT/volt] x Vy [volts] -9667 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64086 [nT]

1989 08 01 04 00

Bx [nT] = -187.0 [nT/volt] x Vx [volts] -190 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -9667 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64086 [nT]

1990 01 30 00 00

Bx [nT] = -187.0 [nT/volt] x Vx [volts] -190 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -9632 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64086 [nT]

1990 12 05 00 00

Bx [nT] = -187.0 [nT/volt] x Vx [volts] -204 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -9590 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64079 [nT]

1992 12 25 00 00

Bx [nT] = -187.0 [nT/volt] x Vx [volts] -698 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -9918 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64109 [nT]

1993 05 04 00 00

Bx [nT] = -187.0 [nT/volt] x Vx [volts] -623 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -9956 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64113 [nT]

1994 01 01 00 00

Bx [nT] = -187.0 [nT/volt] x Vx [volts] -623 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -9956 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64113 [nT]

1994 07 20 00 00
Bx [nT] = -187.0 [nT/volt] x Vx [volts] -591 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -9976 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64113 [nT]

1994 11 22 00 00
Bx [nT] = -187.0 [nT/volt] x Vx [volts] -551 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -9986 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64113 [nT]

1995 02 01 00 00
Bx [nT] = -187.0 [nT/volt] x Vx [volts] -571 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -9957 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64121 [nT]

1995 05 15 00 00
Bx [nT] = -187.0 [nT/volt] x Vx [volts] -549 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -9986 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64113 [nT]

1995 09 01 00 00
Bx [nT] = -187.0 [nT/volt] x Vx [volts] -537 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -9966 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64112 [nT]

1996 01 01 00 00
Bx [nT] = -187.0 [nT/volt] x Vx [volts] -546 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -9957 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64103 [nT]

1996 04 15 00 00
Bx [nT] = -187.0 [nT/volt] x Vx [volts] -521 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -10006 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64104 [nT]

1996 08 15 00 00
Bx [nT] = -187.0 [nT/volt] x Vx [volts] -520 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -9993 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64101 [nT]

1996 11 15 00 00
Bx [nT] = -187.0 [nT/volt] x Vx [volts] -491 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -9990 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64093 [nT]

1997 01 25 00 00
Bx [nT] = -187.0 [nT/volt] x Vx [volts] -491 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -9991 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64093 [nT]

1997 02 15 00 00
Bx [nT] = -187.0 [nT/volt] x Vx [volts] -507 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -9997 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64091 [nT]

1997 06 15 00 00
Bx [nT] = -187.0 [nT/volt] x Vx [volts] -502 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -10014 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64090 [nT]

1997 08 15 00 00
Bx [nT] = -187.0 [nT/volt] x Vx [volts] -488 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -10007 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64092 [nT]

1997 11 15 00 00
Bx [nT] = -187.0 [nT/volt] x Vx [volts] -477 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -9987 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64099 [nT]

1998 01 01 00 00
Bx [nT] = -187.0 [nT/volt] x Vx [volts] -477 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -9990 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64100 [nT]

1998 03 21 00 00
Bx [nT] = -187.0 [nT/volt] x Vx [volts] -487 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -10017 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64100 [nT]

1998 07 23 00 00
Bx [nT] = -187.0 [nT/volt] x Vx [volts] -481 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -10032 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64104 [nT]

1998 10 28 00 00
Bx [nT] = -187.0 [nT/volt] x Vx [volts] -472 [nT]
By [nT] = -201.0 [nT/volt] x Vy [volts] -9996 [nT]
Bz [nT] = -200.0 [nT/volt] x Vz [volts] -64098 [nT]

1989 01 17 00 00
Vx [volts] = 0.004883000 x Dx [dig] -0.0000000
Vy [volts] = 0.004883000 x Dy [dig] -0.0000000
Vz [volts] = 0.004883000 x Dz [dig] -0.0000000

1989 08 01 04 00
Vx [volts] = 0.002441000 x Dx [dig] -0.0073200
Vy [volts] = 0.002441000 x Dy [dig] -0.0048800
Vz [volts] = 0.002441000 x Dz [dig] -0.0073200

1990 01 30 04 00
Vx [volts] = 0.002438000 x Dx [dig] -0.0073100
Vy [volts] = 0.002441000 x Dy [dig] -0.0048800
Vz [volts] = 0.002441000 x Dz [dig] -0.0073200

1990 11 19 08 00
Vx [volts] = 0.002438000 x Dx [dig] -0.0073100
Vy [volts] = 0.002449000 x Dy [dig] -0.0000000
Vz [volts] = 0.002440000 x Dz [dig] -0.0097600

1991 05 03 20 00
Vx [volts] = 0.002469000 x Dx [dig] -0.0000000
Vy [volts] = 0.002469000 x Dy [dig] -0.0000000
Vz [volts] = 0.002461000 x Dz [dig] -0.0000000

1992 12 25 00 00
Vx [volts] = 0.002465000 x Dx [dig] +0.0017250
Vy [volts] = 0.002465000 x Dy [dig] +0.0019720
Vz [volts] = 0.002458000 x Dz [dig] +0.0007370

1993 01 01 00 00
Vx [volts] = 0.002465000 x Dx [dig] +0.0041900
Vy [volts] = 0.002465000 x Dy [dig] -0.0012324
Vz [volts] = 0.002457000 x Dz [dig] -0.0012280

1994 01 01 00 00
Vx [volts] = 0.002465000 x Dx [dig] +0.0041900
Vy [volts] = 0.002465000 x Dy [dig] -0.0020200
Vz [volts] = 0.002457000 x Dz [dig] -0.0020150

1995 01 01 00 00
Vx [volts] = 0.002477000 x Dx [dig] -0.0004500
Vy [volts] = 0.002470000 x Dy [dig] +0.0015800
Vz [volts] = 0.002470000 x Dz [dig] -0.0181800

1995 01 07 00 00
Vx [volts] = 0.002476000 x Dx [dig] +0.0069800
Vy [volts] = 0.002467000 x Dy [dig] +0.0089800
Vz [volts] = 0.002470000 x Dz [dig] -0.0103200

1996 09 26 00 00
Vx [volts] = 0.002480000 x Dx [dig] +0.0051700
Vy [volts] = 0.002470000 x Dy [dig] +0.0080800
Vz [volts] = 0.002470000 x Dz [dig] -0.0078600

1997 12 16 00 00
Vx [volts] = 0.000305512 x Dx [dig] +0.0015627
Vy [volts] = 0.000305206 x Dy [dig] +0.0014833
Vz [volts] = 0.000305176 x Dz [dig] -0.0009583

1998 01 19 00 00
Vx [volts] = 0.000305237 x Dx [dig] -0.0002320
Vy [volts] = 0.000305176 x Dy [dig] +0.0003357
Vz [volts] = 0.000305176 x Dz [dig] -0.0002747

A2 Casey temperature component

1989 01 17 00 00
T[C] = 13.21 [C/volt] x Vt [volts] -5.645 [C]

1992 12 25 00 00
T[C] = 10.00 [C/volt] x Vt [volts] -0.380 [C]

1993 01 01 00 00
T[C] = 10.10 [C/volt] x Vt [volts] +0.300 [C]

1994 01 01 00 00
T[C] = 9.98 [C/volt] x Vt [volts] +0.380 [C]

1995 07 01 00 00
T[C] = 9.99 [C/volt] x Vt [volts] +0.490 [C]

1996 07 01 00 00
T[C] = 10.00 [C/volt] x Vt [volts] +0.390 [C]

1989 01 17 00 00
Vt [volts] = 0.004883000 x Dt [dig] +0.0000000

1992 12 25 00 00
Vt [volts] = 0.000991000 x Dt [dig] +1.9950000

1993 01 01 00 00
Vt [volts] = 0.000991000 x Dt [dig] +1.9979000

1994 01 01 00 00
Vt [volts] = 0.000991000 x Dt [dig] +1.9973000

1995 01 01 00 00
Vt [volts] = 0.002457000 x Dt [dig] +1.8161000

1995 01 07 00 00
Vt [volts] = 0.002449000 x Dt [dig] +1.8602000

1996 09 26 00 00
Vt [volts] = 0.002460000 x Dt [dig] +1.8210000

1997 12 16 00 00
Vt [volts] = 0.000305451 x Dt [dig] -0.0001680

1998 01 19 00 00
Vt [volts] = 0.000306617 x Dt [dig] +0.0011713

A3 Davis X, Y and Z components

1988 02 16 10 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3600 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16000 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52150 [nT]

1989 12 01 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3653 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16042 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52141 [nT]

1990 03 01 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3596 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16000 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52158 [nT]

1990 11 01 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3644 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16014 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52151 [nT]

1991 01 01 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3557 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16042 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52142 [nT]

1992 01 08 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3632 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16074 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52121 [nT]

1992 03 01 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3611 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16051 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52128 [nT]

1992 04 01 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3597 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16077 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52142 [nT]

1992 11 16 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3622 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16069 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52142 [nT]

1993 01 01 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3738 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16043 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52129 [nT]

1993 04 10 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3691 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16034 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52145 [nT]

1993 10 27 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3732 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16074 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52174 [nT]

1994 01 01 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3728 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16077 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52171 [nT]

1994 02 15 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3714 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16065 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52169 [nT]

1994 03 15 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3690 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16060 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52174 [nT]

1994 04 10 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3675 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16058 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52176 [nT]

1994 10 02 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3724 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16080 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52171 [nT]

1995 01 01 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3729 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16074 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52170 [nT]

1995 02 10 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3702 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16065 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52172 [nT]

1995 04 01 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3689 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16030 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52190 [nT]

1995 08 10 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3694 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16019 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52160 [nT]

1995 11 01 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3720 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16033 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52150 [nT]

1996 02 01 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3745 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16040 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52130 [nT]

1996 04 06 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3714 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16041 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52137 [nT]

1996 06 10 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3712 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16043 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52139 [nT]

1996 07 07 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3703 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16043 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52142 [nT]

1996 09 01 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3717 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16042 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52150 [nT]

1996 11 07 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3741 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16048 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52138 [nT]

1996 12 07 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3756 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16054 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52127 [nT]

1997 01 01 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3748 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16061 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52136 [nT]

1997 02 09 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3735 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16062 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52150 [nT]

1997 03 05 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3724 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16055 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52161 [nT]

1997 04 08 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3700 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16053 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52168 [nT]

1997 11 03 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3736 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16062 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52169 [nT]

1997 12 11 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3754 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16077 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52164 [nT]

1998 01 01 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3746 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16078 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52158 [nT]

1998 02 19 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3737 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16077 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52159 [nT]

1998 03 09 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3720 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16071 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52164 [nT]

1998 04 18 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3695 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16068 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52172 [nT]

1998 10 16 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3722 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16072 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52170 [nT]

1998 11 16 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3739 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16102 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52148 [nT]

1999 01 01 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3747 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16100 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52157 [nT]

1999 02 19 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3698 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16068 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52154 [nT]

1999 03 17 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3720 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16071 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52164 [nT]

1999 04 14 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3702 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16073 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52157 [nT]

1999 09 10 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3698 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16070 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52165 [nT]

1999 10 08 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3715 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16075 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52158 [nT]

1999 12 08 00 00
Bx [nT] = -194.8 [nT/volt] x Vx [volts] +3737 [nT]
By [nT] = +200.7 [nT/volt] x Vy [volts] -16088 [nT]
Bz [nT] = +195.6 [nT/volt] x Vz [volts] -52146 [nT]

1998 02 16 10 00
Vx [volts] = 0.000976000 x Dx [dig] -0.0009760
Vy [volts] = 0.000976000 x Dy [dig] -0.0029300
Vz [volts] = 0.000976000 x Dz [dig] -0.0029300

1988 03 16 07 45
Vx [volts] = 0.002439000 x Dx [dig] -0.0024390
Vy [volts] = 0.002439000 x Dy [dig] -0.0097600
Vz [volts] = 0.002439000 x Dz [dig] -0.0000000

1989 01 01 00 00
Vx [volts] = 0.002459000 x Dx [dig] -0.0049710
Vy [volts] = 0.002437000 x Dy [dig] -0.0146200
Vz [volts] = 0.002464000 x Dz [dig] -0.0024600

1989 12 01 00 00
Vx [volts] = 0.002459000 x Dx [dig] -0.0049170
Vy [volts] = 0.002437000 x Dy [dig] -0.0194900
Vz [volts] = 0.002464000 x Dz [dig] -0.0049300

1991 03 24 00 00
Vx [volts] = 0.002438000 x Dx [dig] -0.0007320
Vy [volts] = 0.002439000 x Dy [dig] +0.0012200
Vz [volts] = 0.002439000 x Dz [dig] -0.0004880

1992 03 05 00 00
Vx [volts] = 0.002435000 x Dx [dig] -0.0034100
Vy [volts] = 0.002434000 x Dy [dig] +0.0007300
Vz [volts] = 0.002435000 x Dz [dig] +0.0012200

1992 12 11 07 15
Vx [volts] = 0.002436000 x Dx [dig] -0.0042230
Vy [volts] = 0.002436000 x Dy [dig] +0.0016500
Vz [volts] = 0.002436000 x Dz [dig] +0.0020300

1994 04 07 00 00
Vx [volts] = 0.002435000 x Dx [dig] -0.0037990
Vy [volts] = 0.002435000 x Dy [dig] -0.0013640
Vz [volts] = 0.002435000 x Dz [dig] -0.0010720

1995 01 01 00 00
Vx [volts] = 0.002457000 x Dx [dig] -0.0164630
Vy [volts] = 0.002467000 x Dy [dig] -0.0002470
Vz [volts] = 0.002436000 x Dz [dig] -0.0019490

1995 07 03 00 00
Vx [volts] = 0.002457000 x Dx [dig] +1.2271480
Vy [volts] = 0.002467000 x Dy [dig] -0.0002470
Vz [volts] = 0.002436000 x Dz [dig] +1.7078340

1996 01 01 00 00
Vx [volts] = 0.002459000 x Dx [dig] +1.2338000
Vy [volts] = 0.002469000 x Dy [dig] +0.0009880
Vz [volts] = 0.002438000 x Dz [dig] +1.7066930

1997 02 07 00 00
Vx [volts] = 0.002459480 x Dx [dig] +1.2376104
Vy [volts] = 0.002469319 x Dy [dig] +0.0029632
Vz [volts] = 0.002438846 x Dz [dig] +1.7135332

1997 12 31 00 00
Vx [volts] = 0.000307342 x Dx [dig] +0.0016013

Vy [volts] = 0.000327955 x Dy [dig] +0.3507477
Vz [volts] = 0.000304776 x Dz [dig] +0.0019841

1998 01 01 00 00
Vx [volts] = 0.000307342 x Dx [dig] +0.0016013
Vy [volts] = 0.000308547 x Dy [dig] -0.0042271
Vz [volts] = 0.000304776 x Dz [dig] +0.0019841

1999 02 25 00 00
Vx [volts] = 0.000307815 x Dx [dig] +0.0013452
Vy [volts] = 0.000308642 x Dy [dig] -0.0059259
Vz [volts] = 0.000304785 x Dz [dig] +0.0010850

A4 Davis temperature component

1988 02 16 10 00
T[C] = 10.19 [C/volt] x Vt [volts] -0.409 [C]

1992 05 21 00 00
T[C] = 10.20 [C/volt] x Vt [volts] -0.470 [C]

1995 01 01 00 00
T[C] = 10.03 [C/volt] x Vt [volts] -0.070 [C]

1988 02 16 10 00
Vt [volts] = 0.000976000 x Dt [dig] +1.9816500

1990 09 17 00 00
Vt [volts] = 0.000982000 x Dt [dig] +1.9981640

1991 03 17 00 00
Vt [volts] = 0.001011000 x Dt [dig] +2.0027300

1992 03 05 07 30
Vt [volts] = 0.001011000 x Dt [dig] +1.9932750

1992 12 11 08 00
Vt [volts] = 0.001011000 x Dt [dig] +1.9934280

1994 04 07 09 00
Vt [volts] = 0.001011000 x Dt [dig] +1.9959570

1995 01 01 00 00
Vt [volts] = 0.002438000 x Dt [dig] +1.9660620

1996 01 01 00 00
Vt [volts] = 0.002440000 x Dt [dig] +1.9695530

1997 02 07 00 00
Vt [volts] = 0.002440810 x Dt [dig] +1.9714425

1997 12 31 00 00
Vt [volts] = 0.000304971 x Dt [dig] +0.0001220

APPENDIX B
SAMPLE OUTPUT OF A PROCESSED FILE

The following is an example of a single days magnetometer data processed to an ASCII format (known as the ASP format). The first record is the file name and in this example it identifies it as processed Davis 1999 (day 339) magnetic data. The second record specifies the date at which the ASCII file was created. Record three is the offset baseline values for X, Y, and Z. The next 8640 records are data values in the order X, Y and Z. The remaining records are the calibration equations used for that days data. To uncompress the data to a [nT] value use:

value (nT) = (offset * 1000 + data value)/10.0

blank 999999999999999 should be set to 99999.9 for each X, Y and Z entry.

```
PDAMG99.339
 Created on Thu Mar 23 18:10:01 2000
   27 -166 -522
 999999999999999
 999999999999999
 999999999999999
 999999999999999
 999999999999999
 999999999999999
 999999999999999
  6423 1566 4479
  6422 1573 4497
  6428 1577 4515
  6439 1580 4533
         .
         .
         .
         .
  6126  940 4694
  6117  941 4705
  6111  943 4722
  6108  955 4752
#
1) CALIBRATION EQUATION(S):
CALIBRATION EQUATION RECORDS
NUMBER OF RECORDS: 1
1999 281  0  0
Bx [nT] = (-0.05996)[nT/Dig]+(  3714.74)[nT]
By [nT] = ( 0.06194)[nT/Dig]+(-16076.19)[nT]
Bz [nT] = ( 0.05962)[nT/Dig]+(-52157.79)[nT]
T [C]   = ( 0.00979)[ T/Dig]+(    19.81) [C]
```

APPENDIX C
ABSOLUTES BASELINE DATA

The following has been extracted from MAGINFO.CSY and MAGINFO.DAV. These files are used by the AGSO programs REDUCE.EXE and COMPARE.EXE in the reduction of magnetic absolutes. The data below consists of the AGSO form of the baseline values and scale values together with relevant range and drift information for the X, Y, Z, F and T components of the variometer and absolute instruments. The application of this data to magnetic absolute reductions is discussed in Crosthwaite (1994).

C1 Casey

```
# Temperature calibration data for 25-Jan -97 back dated to 1-Jan-97.

VARIOMETER Cas EDA 1997 000.0000 1997 025.9999\
  A\
  CF\
   -538.3  -0.4630    0.0       0.0      0.0 \
  -9968.0   0.0      -0.4959    0.0      0.0 \
 -64110.0   0.0       0.0      -0.4940   0.0 \
    0.00    1.000                       0.0 \
   18.60    0.0246                     20.00\
    0  0  0  0  0\
    0  0  0  \
    0.00   0.00     0.00  0.025\
    F
# Analogue Conditioner calibration performed 25-Jan-97 (Day 025).

VARIOMETER Cas EDA 1997 026.0000 1997 349.9999\
  A\
  CF\
   -492.0  -0.4630    0.0       0.0      0.0 \
  -9991.6   0.0      -0.4964    0.0      0.0 \
 -64091.4   0.0       0.0      -0.4942   0.0 \
    0.00    1.000                       0.0 \
   18.60    0.0246                     20.00\
    0  0  0  0  0\
    0  0  0  \
    0.00   0.00     0.00  0.025\
    F
```

```
# Analogue Conditioner calibration 16-Dec-97 (Day 350).
VARIOMETER Cas EDA 1997 350.0000 1998 022.9999\
  A\
  CF\
   -491.3  -0.0571    0.0       0.0      0.0 \
  -9990.3   0.0      -0.0613    0.0      0.0 \
 -64092.8   0.0       0.0      -0.0610   0.0 \
    0.00    1.000                      0.0 \
    0.40    0.0031                     20.00\
    0  0  0  0  0\
    0  0  0  \
    0.00   0.00     0.00  0.025\
    F

# Analogue Conditioner calibration 19-Jan-98 (Day 019).
# Baseline calibration 22-Jan-98 (Day 022).
# Results applied from 23-Jan-98 (Day 023).
VARIOMETER Cas EDA 1998 023.0000 1999 021.9999\
  A\
  CF\
   -477.0  -0.0571    0.0       0.0      0.0 \
  -9987.1   0.0      -0.0613    0.0      0.0 \
 -64099.3   0.0       0.0      -0.0610   0.0 \
    0.00    1.000                      0.0 \
    0.40    0.0031                     20.00\
    0  0  0  0  0\
    0  0  0  \
    0.00   0.00     0.00  0.025\
    F
```

```
# Analogue Conditioner calibration 21-Jan-99 (Day 021).
VARIOMETER CSY CSY 1999 022.0000 2010 365.9999\
  A\
  CF\
   -475.9  -0.0572    0.0       0.0      0.0 \
  -9986.5   0.0      -0.0614    0.0      0.0 \
 -64099.3   0.0       0.0      -0.0610   0.0 \
    0.00    1.000                      0.0 \
    0.40    0.0031                     20.00\
    0  0   0  0  0\
    0  0   0  \
    0.00    0.00     0.00  0.025\
    F
```

C2 Davis

```
# 04 March 1996 (00:00:00 UT) to 31 December 1996 (00:00:59 UT).
# Baseline data from a calibration set performed in January 1995.
# AGSO scale values from an analogue conditioner calibration of 04
March 1996.
VARIOMETER Dav EDA 1996 4.00000 1996 365.99999\
        C\
        CF\
         3484    -0.4790    0.0       0.0       0.0\
      -16080     0.00       0.4950    0.0       0.0\
      -51837     0.00       0.0       0.4770    0.0\
           0.00    1.0000                       0.0\
          19.62    0.0249                      20.00\
           0  0   0  0 0\
           0  0   0  \
         0.00  -0.00     -0.00  0.025\
        C
```

```
# 01 January 1997 (00:00:00 UT) to 30 January 1998 (09:29:59 UT).
# An analogue conditioner calibration of 07 February 1997 was
# never included in MAGINFO.DAV. It has NOT been retrospectively
# applied as 1997 has already been processed to WDC-A standards.
VARIOMETER Dav EDA 1997 0.00000 1998 30.39582\
        C\
        CF\
         3484    -0.4790    0.0       0.0       0.0\
      -16080     0.00       0.4950    0.0       0.0\
      -51837     0.00       0.0       0.4770    0.0\
          0.00   1.0000                         0.0\
         19.62   0.0249                        20.00\
          0  0   0  0 0\
          0  0   0  \
        0.00  -0.00     -0.00  0.025\
        C

# 30 January 1998 (09:30:00 UT) to 25 February 1999 (02:30:00 UT).
# Constants modified due to an analogue conditioner calibration.
VARIOMETER Dav EDA 1998 30.39583 1999 56.10417\
        C\
        CF\
         3723.7  -0.0599    0.0       0.0       0.0\
      -16080.8   0.0        0.0619    0.0       0.0\
      -52170.6   0.0        0.0       0.0596    0.0\
          0.00   1.0000                         0.0\
         -0.47   0.0031                        20.00\
          0  0   0  0 0\
          0  0   0  \
        0.00  -0.00     -0.00  0.025\
        C
```

```
# 25 February 1999 (02:30:01 UT) to present time.
# Constants modified due to an analogue conditioner calibration.
VARIOMETER Dav EDA 1999 56.10418 2010 365.99999\
        C\
        CF\
         3723.7  -0.0600    0.0       0.0       0.0\
       -16081.2   0.0       0.0619    0.0       0.0\
       -52170.8   0.0       0.0       0.0596    0.0\
           0.00   1.0000                          0.0\
          -0.47   0.0031                        20.00\
          0  0   0  0 0\
          0  0   0  \
        0.00  -0.00     -0.00  0.025\
        C
```

APPENDIX D
OBSERVATION ERRORS

The magnetic field intensity of the X, Y and Z components of the Earth's magnetic field can be expressed in terms of Total field F, Declination D, and Inclination I, as follows:

Bx = F * COS(I) * COS(D)	nT	(1)
By = F * COS(I) * SIN(D)	nT	(2)
Bz = F * SIN (I)	nT	(3)

The error in Bx, By and Bz can be expressed as the partial differentials of Equations (1), (2) and (3) as follows:

∂Bx = -F * COS(I) * SIN (D) * ∂D	nT	(4)
∂By = F * COS(I) * COS(D) * ∂D	nT	(5)
∂Bx = -F * COS (D) * SIN(I) * ∂I	nT	(6)
∂By = -F * SIN (D) * SIN(I) * ∂I	nT	(7)
∂Bz = F * COS (I) * ∂I	nT	(8)

Typical F, D and I baseline values (based on 1999 data) for the stations are as follows:

Casey:
F = 64885 nT
D = -93 degrees
I = -81 degrees

Davis:
F = 54710 nT
D = -77 degrees
I = -72 degrees

Mawson:
F = 49480 nT
D = -65 degrees
I = -68 degrees

Macquarie Island:
F = 64500 nT
D = 30 degrees
I = -79 degrees

Applying the above F, D and I values to Equations (4) through (8) produces the error data shown in the Tables 2 and 3.

Table 2. *Affect of declination error on the X, Y and Z field components.*

		Casey			Davis			Mawson			Macquarie Island		
∂D (')	∂D (°)	∂Bx (nT)	∂By (nT)	∂Bz (nT)	∂Bx (nT)	∂By (nT)	∂Bz (nT)	∂Bx (nT)	∂By (nT)	∂Bz (nT)	∂Bx (nT)	∂By (nT)	∂Bz (nT)
0.1	0.0017	0.3	0.0	0.0	0.5	0.1	0.0	0.5	0.2	0.0	-0.2	0.3	0.0
1.0	0.0167	2.9	-0.2	0.0	4.8	1.1	0.0	4.9	2.3	0.0	-1.8	3.1	0.0
6.0	0.1000	17.7	-0.9	0.0	28.8	6.6	0.0	29.3	13.7	0.0	-10.7	18.6	0.0

Table 3. *Affect of inclination error on the X, Y and Z field components.*

		Casey			Davis			Mawson			Macquarie Island		
∂D (')	∂D (°)	∂Bx (nT)	∂By (nT)	∂Bz (nT)	∂Bx (nT)	∂By (nT)	∂Bz (nT)	∂Bx (nT)	∂By (nT)	∂Bz (nT)	∂Bx (nT)	∂By (nT)	∂Bz (nT)
0.1	0.0017	-0.1	-1.9	0.3	0.3	-1.5	0.5	0.6	-1.2	0.5	1.6	0.9	0.4
1.0	0.0167	-1.0	-18.6	3.0	3.4	-14.7	4.9	5.6	-12.1	5.4	16.0	9.2	3.6
6.0	0.1000	-5.9	-111.7	17.7	20.4	-88.5	29.5	33.8	-72.6	32.4	95.7	55.3	21.5

APPENDIX E
DATA PROCESSING PROCEDURE

The following is an outline of the procedure used to process Casey and Davis magnetic data from the raw binary format to a the World Data Centre A (WDC-A) format. For convenience Casey 1997 is used as an example in some commands shown throughout the text.

E1 Using PROCESS (produces ASP format files)

E1.1 From Marvin

Type SD ADAS:[PROCESS]. Sets the correct directory.
Type IDL. Puts you in the IDL directory.
Type @WDCRUN. Executes the IDL processing software (use if generating data plots).

Or

Type @WDCRUN2. Executes the IDL processing software (use if you do not wish to generate data plots).

E1.2 From the IDL application Widget

Select CMAGCAL.DAT
Select CLOGCAL.DAT
Select CBLANK97.DAT
Select Path: ADAS:[1997.CASEY]. Sets the correct directory.
Select Enter

E1.3 To process and plot a months data

Select CASMAG97.001. Selects the first mag file to be processed.
Select CASMAG97.031. Selects the last mag file to be processed.

E1.4 Expected output

Postscript plot files are produced. (eg. PS.001) which need to be deleted manually. Processed data files are produced. (eg PCAMG97.001) with 8640 data records.

E2 Producing WDC-A data format files

E2.1 To produce 1-minute files from Marvin

Type SD RDLCAL. Sets the correct directory.
Type RUN WDC. Executes WDC.EXE (Fortran).
Type 1. For minute files.
Select ENTER.
PATH TO INPUT DATA: ADAS:[1997.CASEY]. Sets the correct directory.
PATH TO OUTPUT DATA: ADAS:[1997.CASEY]. Sets the correct directory.
INPUT START FILE: PCAMG97.001.
INPUT LAST FILE: PCAMG97.365.

Wait while program processes information. Messages will be displayed to indicate output files of the name WCAMG9701.MIN to WCAMG9712.MIN (one for each month) are produced.

E2.2 To Remove unwanted characters

Rename the 1-minute files produced in Step 2.1 so DOS can handle the file name length (e.g. from WCAMG9701.MIN to CAMG9701.MIN). Transfer renamed files from the VAX to a PC (16 bit system only).

From the PC (16 bit system only):
Run PCT.EXE.
Select the appropriate file.
Select w.
Press any key but ESC.
Select F2.
Press any key.
Press ESC.
Repeat for the 12 month files for the year.

E2.3 To fix problems with the format

From the PC (16 bit system only):
Modify CSY.DAT or DVS.DAT to reflect appropriate year.
Type MVFIX_AU < CSY.DAT wait a few minutes while it processes a years worth of files.

E2.4 To create a data archive

Type LHA A CSY97_1M.LZH CSY9701M.XYZ. An archive will be created and first months data zipped. Repeat until first 6 months data is zipped.

Then repeat using LHA A CSY97_2M.LZH CSY9707M.XYZ. Repeat until second 6 months data is zipped.

APPENDIX F
DOCUMENT LIST

F1 AAD publications

Burns, G.B. and Klekociuk, A.R. (1994). Davis and Casey Magnetic Field Data. In: Morris, R.J. (Ed.). *ANARE Research Notes 92*. Pp. 17–41.

Paptashvili, V.O., Belov, B.A., Burtsev, Yu. A., Pimenov, I.A., Burns, G.B., Morris, R.J., Neudegg, D.A. and Symons, L.P. (1996). Comparison of High Resolution Quartz and Fluxgate Magnetometer Data Recorded at Davis, Antarctica. In: Morris, R.J. (Ed.). *ANARE Research Notes 95*. Pp. 81–92.

F2 AAD Reports on data collection and processing

Dymond, M. (1990). Analogue Conditioner Rack Technical Manual.

Giles, A.B. (1991). An Integrated Archive Data System for UAP.

Giles, A.B. (1992). ASP Magnetometer Routine Data Processing System.

Giles, A.B. (1992). Creating World Data Centre(WDC) Format Files.

Newbery, S.M. (1993). (Modification of) Creating World Data Centre (WDC) Format Files.

Symons, L.P. (1997). Users Guide To VIEW Version 5.00.

Symons, L.P. (1997). Users Guide To ADAS Version 4.0.

Roberts, P.I. (2000). Macquarie Island Atmospheric and Space Physics 1999 Annual Report.

F3 AGSO Publications and internal reports

Crosthwaite, P. (1994). Personal Computer Data Acquisition (MACQ) and Reduction Software for Geomagnetic Observatories: Reference Manual. *Geomagnetism Note 1994-12*.

Dennis, S. (1995). Casey Baselines.

F4 Casey magnetometer internal AAD reports

Breed, A. and Schneider, D. (1996). Casey Fluxgate Magnetometer Annual Report.

Breed, A. (1998). Casey Fluxgate 1998 Annual Report.

Dubovinsky, M. and Langmaid, C. (1994). Casey Magnetometer Calibration - January 1995.

Dubovinsky, M. and Langmaid, C. (1995). Casey Magnetometer Calibration - January 1995.

Hesse, M. and Roberts, P.I. (1991). Casey Magnetometer Calibration October/November 1991.

Langmaid, C. and Schneider, D. (1996). Fluxgate Magnetometer Annual Calibration Report. Casey Station Antarctic.

McLoughlin, R. (1990). Casey Magnetometer Calibration 20/07/90.

McLoughlin, R. (1990). Casey Magnetometer Calibration 05/02/89 - 25/05/90.

Symons, L.P. (1993). Casey Magnetometer Calibration Feb 1993.

Symons, L.P. (1993). Casey Fluxgate Magnetometer Baseline Determination.

Symons, L.P. (1993). Routine Data Logging. Description of Data Casey 1993.

Symons, L.P. and Dubovinsky, M. (1994). Casey Magnetometer Calibration January 1994.

F5 Davis magnetometer internal AAD reports

Hesse, M. (1991). Davis Magnetometer Calibration 1990.

Hesse, M. (1991). Routine Data Logging. Description of Data Davis 1990.

Hesse, M. and Symons, L.P. (1991). Davis Magnetometer Calibration September 1991.

Hesse, M. and Symons, L.P. (1992). Davis Magnetometer Calibration January 1992.

Hesse, M., Symons, L.P. and Burns, G.B. (1992). Davis Magnetometer Calibration March 1992.

Leschinski, R. (1997). Davis Magnetometer Calibration - April 1997.

Neudegg, D. (1992). 1992 Davis Magnetometer Calibration.

Parcell, S. (1994). Fluxgate Magnetometer Annual Report and Calibration Davis 1993.

Symons, L.P. (1995). 1994 Davis Magnetometer Calibration Report.

Symons, L.P. (1996). Davis Magnetometer Calibration - May 1996.

Roberts, P.I. (1999). Davis Magnetometer: Adopted Baselines for 1998.

F6 Station monthly reports

Various authors. Casey Fluxgate Magnetometer Monthly Reports. Archived, in text format, on the Australian Antarctic Division's VAX computer in the directory PHYS:[CASEY.REPORTS.YYYY].

Various authors. Davis Fluxgate Magnetometer Monthly Reports. Archived, in text format, on the Australian Antarctic Division's VAX computer in the directory PHYS:[DAVIS.REPORTS.YYYY].

31. ANTARCTIC MIDDLE AND UPPER ATMOSPHERE PHYSICS STRATEGIC PLAN 2001-05

R.J Morris and D.J. Murphy

Atmospheric and Space Physics
Australian Antarctic Division
Kingston Tasmania 7050 Australia
(email: ray.morris@aad.gov.au)

Abstract

This paper presents a strategic plan for Australian middle and upper atmospheric physics for the 2001–05 interval, to be conducted at the Antarctic stations, Davis (68.6°S, 78.0°E), Mawson (67.6°S, 62.9°E) and Casey (66.3°S, 110.5°E), and the sub-Antarctic station at Macquarie Island (54.5°S, 159°E). The strategic plan addresses the four Government goals for the Australian Antarctic program: to maintain the Antarctic Treaty System and enhance Australia's influence within the system; to protect the Antarctic environment; to understand the role of Antarctica in the global climate system; and to undertake scientific work of practical, economic and national significance. For each goal we discuss 'key scientific outputs' and associated 'assessment milestones'. An abbreviated version of this strategic plan is located at the Australian Antarctic Division web site – http://www.aad.gov.au/. The strategic plan is dynamic and will undergo annual adjustments.

31.1 Introduction

Antarctic atmosphere studies are vital to understanding anthropogenic and natural climate change, and space weather. Global warming via anthropogenic influence, 'the greenhouse effect', is the dominant global climate issue. Greenhouse warming in the troposphere (0–10 km) is associated with enhanced cooling in the upper atmosphere (stratosphere: 10-50 km, mesosphere: 50-90 km, and thermosphere: above 90 km). However, recent evidence suggests that the upper atmosphere is cooling faster than is predicted by present models, implying either other processes of change are operating or that we do not yet understand fully the pathways of greenhouse warming. The variation in climate trends through the altitude ranges of the atmosphere, and across the seasons, provides evidence of the processes via which changes are occurring (e.g. carbon dioxide and/or methane increase; ozone destruction; natural solar variability; aerosols). The Antarctic upper atmosphere has demonstrated that it can change rapidly in a manner detrimental to ecological systems. The factors that resulted in the ozone hole being predominantly an Antarctic phenomena, global extreme cold temperatures and the existence of cloud particles at high altitudes

allowing heterogeneous catalytic reactions, give additional emphasis to the need for upper atmosphere climate trend studies in Antarctica.

Solar energy absorbed by the Earth's atmosphere changes the temperature of our environment and the atmospheric motions as well. These motions, winds and atmospheric waves, transport the sun's energy over the Earth's surface into the middle and upper atmosphere, bringing about unexpected changes in the structure of the Earth's atmosphere. The temperature in the stratosphere increases with height because the ozone layer is absorbing energy as it stops UV radiation from reaching the Earth's surface. And yet the warmest part of the polar winter stratosphere is up to 15 km above the height of maximum heating by UV radiation. Similarly, the coldest part of Earth's atmosphere is the summer mesopause, some 90 km above the Earth's surface, is bathed in continuous sunlight. The reason for both these apparent anomalies has been traced to atmospheric motions, called atmospheric gravity waves (AGW's), which vary on time scales of less than a day.

AGW's are generally produced by topography and weather systems that surround us in the troposphere. They transport momentum upward into the stratosphere, mesosphere and lower thermosphere. As they move upward they grow in amplitude due to decreasing density with height. Eventually, they grow to a point where they become unstable, 'break' and deposit their momentum. This generates forces that drive horizontal and vertical air circulations. The downward transport of stratospheric air explains, the warm polar winter stratopause and summer air moving up into the mesosphere, creates the intensely cold summer polar mesopause. The impact of gravity wave breaking at one level has profound effects on the atmosphere at much lower levels – 'downward control'. In such ways, AGW's modify and drive the transport and mixing of much of the middle atmosphere.

Other atmospheric motions, such as tides, which have periods of 24, 12 and 8 hours, or planetary waves, which oscillate over a number of days, have the potential to divert a gravity wave's path or to cause it to break before it has reached mesospheric heights. These interactions affect both the tides and the planetary waves as well as AGW's. A single planetary wave or tide can cover a significant proportion of the globe making it necessary to observe at different longitudes and observations in Antarctica are vital to global investigations. In addition, the atmosphere is changing from minute-to-minute, day-to-day and season-to-season and all these changes affect the state of the atmosphere. As the middle and upper atmosphere role in the absorption and distribution of solar energy becomes clearer, the need for further measurements of this region becomes more apparent. If the global nature of atmospheric processes is to be characterised, then observations at polar latitudes, where data are sparse, are needed.

Energy from the more extreme parts of the solar radiation spectrum does not reach the lower atmosphere. On the Earth's surface we are protected by exchanges that are so violent, that the atoms and molecules in

the upper atmosphere are stripped of some of their electrons. Known as the ionosphere, this region reflects the radio waves used in over-the-horizon-radars (OTHR) and other HF systems. Furthermore magnetic and electric forces affect the ionised atmosphere as it responds to solar heating. A feature of these processes is the global-scale thermospheric wind pattern, which is driven by solar heating and ionospheric motions resulting from collisions between ions and neutrals. In addition, at high-latitudes, where the magnetic field lines dip into the Earth, electric fields developed in the magnetosphere set up ionospheric convection flows that alter the thermospheric motion further. Energy from the sun is also transmitted to the Earth through magnetic interactions between the solar wind and Earth's magnetic field. The resulting changes affect the thermospheric circulation in the polar regions and accelerate particles back at Earth forming the aurora. The magnetosphere located above the ionosphere and extending out to about 10 Earth radii on the dayside, is a tenuous ionised plasma region populated predominantly with protons and electrons under the control of the Earth's magnetic field. This region also includes the more energetic particles in the Van Allen radiation belts. It is precipitation of these particles that contribute to the northern and southern aurora.

High-latitude processes is the key to understanding the coupling between the thermosphere, ionosphere and magnetosphere and to improving the forecasting of potentially destructive changes in the solar wind – 'space weather'. Measurements of the motions of the ionosphere and thermosphere at high-latitudes improve our understanding of the underlying physical processes and allow their inclusion in predictive computer models. This in turn aids forecasting the geospace environment changes and their impact on modern systems such as satellites, GPS and OTHR, and the development of mitigation strategies that can prevent these systems from being rendered inoperable.

Currently, only the low-altitude consequences of the redistribution of global energy by the greenhouse effect are considered in the climate change program. Since the middle and upper atmosphere may also be affected, this priority area should be expanded to include the geospace environment out to a distance of 100 000 km. It is important to understand the energy input processes from the solar wind into the magnetosphere, and down though the ionosphere to the lower atmosphere. Aspects of these processes require long term synoptic measurements over one or more sunspot cycles (11 years) or the recovery period of the ozone layer.

Australia has claim to over 40% of the Antarctic continent and currently operates three stations on the coastal edge and one island station. Both politics and science have motivated this occupation in the past. The year-round occupation of these stations is driven by the need for synoptic observations of the atmospheric, ionospheric and geospace environments, and their contribution to national and international research. Australia's influence in the Antarctic Treaty System (ATS) is

related to the quality of the data collected in the AAT and the scientific investigations of that data.

It is important to integrate the government's goals for our national Antarctic science program, with international objectives. The stature of Australian Antarctic science will increase through a commitment to international collaboration. Timely distribution of our data to national and international end-users enhances our standing in international fora and contributes to the identification of regional and global processes. Satellites, which do not provide the temporal coverage necessary to parameterise short timescale phenomena, can be used in conjunction with our measurements to produce potent data sets.

31.1.1 Significant questions on the dynamics and thermodynamics of the middle atmosphere include the following

- What is the present day structure and circulation of the middle atmosphere, and what are the key physical and chemical processes that control this region?

- What is the climatology of the stratosphere and mesosphere regions of the Antarctic atmosphere?

- What is the structure of the stratospheric ozone layer above Davis and how does it respond to stratospheric dynamics?

- What are the annual and decadal trends in temperature, winds and densities throughout the stratosphere and mesosphere, and can they provide evidence of global climate change, and the rate of any change?

- What factors control the formation of polar stratospheric and mesospheric clouds, and are they related to global change and the impact of human endeavours on the polar atmosphere?

- How is momentum transferred from the lower atmosphere to the mesosphere and, in particular, what role does turbulence play?

- What evidence is there of a sun-weather coupling mechanism via the near Earth atmospheric electric field?

31.1.2 Significant questions on upper atmosphere and space weather include

- What ionospheric and magnetospheric parameters are needed to improve space weather prediction and nowcasting?

- What processes and regions are involved in the formation of polar patches and associated irregularities and how do they change as they convect across the polar cap?

- What factors apart from the interplanetary magnetic field (IMF), control ionospheric convection? How does ionospheric convection vary with height and when variations occur, what are the causes?

- Is there any further evidence linking thermospheric vertical winds to the cross-polar jet? How well do models predict the wind patterns during active solar times? Can the models be improved?

- What direct evidence of polar cap heating by gravity waves exists in observations from Mawson and Davis? What causes upwellings in the thermospheric wind poleward of the auroral zone?

- What determines which Travelling Ionospheric Disturbances (TID's) generated in the auroral zone travel all the way to the equator, affecting HF communication systems and OTHR's such as JORN?

- What is the location of the open/closed geomagnetic field line boundary and its dynamics at high latitudes?

- How can simple ground-based observations of magnetic signatures contribute to magnetosphere and ionosphere diagnostics?

- What is the plasmapause variability role in magnetospheric/ ionospheric convection?

31.2 Goal 1 – To maintain the Antarctic Treaty system and enhance Australia's influence within the system

Data provision and collaborative involvement in international programs enhance Australia's reputation as a science-faring nation and gives greater credence to its presence in the Antarctic Treaty System (ATS). The data supplied by Australia, regardless of the experiments, should be of high quality and well calibrated. World Data Centres and international collaborative groups seek these data. Three groups include the TIMED satellite mission and two SCOSTEP programs, S-RAMP (STEP, Results, Analysis and Modelling Program) and PSMOS (Planetary Scale Mesopause Observing System). Membership of the Scientific Committee on Antarctic Research (SCAR) Solar Terrestrial and Astrophysical Research (STAR) working group ensures Australia's representation in this international forum. Working groups and projects set up by IAGA (International Association of Geomagnetism and Aeronomy) and IUGG (International Union of Geodesy and Geophysics) set the agenda for much of the STEP (Solar Terrestrial Energy Program) science carried out in Australia, and those conducted by GEM (Geospace Environment and Modelling), CEDAR (Coupling Energy and Dynamics of Atmospheric Regions), CLUSTER, FedSat and others.

Atmospheric Science needs permanent stations throughout the year and provides a justification for the operation of Australia's southern stations. The ATS blossomed through the conduct of meteorology and upper atmospheric sciences during the IGY (1957-1958). An increasing number of participants continue to extend the spatial and temporal coverage of the Antarctic. More sophisticated equipment and improved scientific methodology permit more powerful studies of this region. This has happened because the polar region is uniquely important for ground-based studies of the middle atmosphere and geospace. Indeed at the SCAR 1994 meeting in Rome - SCAR noted:

- the importance of achieving a better understanding of global change;
- the importance of long term data bases in evaluating global change;
- the value of the polar regions for studies of possible solar-terrestrial modulation of climate and weather;
- and that there has been substantial recent progress in identifying possible mechanisms for this coupling;

and urges;

- international organisations and national Antarctic operators to give strong support for long term monitoring of important solar-terrestrial environmental variables in Antarctica, and recommends;
- continual study of these data, and support for development of new instruments and theoretical models.

31.2.1 Key scientific output 1.1 – Continual collection and quality assurance of solar-terrestrial data sets

Continue the collection and quality assurance of data sets. The transfer of these data sets is to be expedited by the installation of data servers that will automate transfer to Kingston and designated web sites. Contribute to NDSC (Network for Detecting Stratospheric Change) as an important group with links to Montreal Protocol monitoring.

31.2.2 Assessment milestones 1.1

- That Australia continues to implement the SCAR 1994 resolution on the solar-terrestrial environment.
- Atmospheric science continues to provide justification for permanent Antarctic stations.
- Continue to take advantage of the unique geographic location of Antarctica to increase scientific understanding of the physics of geospace, the structure of the geomagnetic field, and the physics and chemistry of the upper, middle and lower atmosphere.
- Automated transfer of data to Kingston and to designated web sites.
- Play an active role in collaborative programs such as the ground-based 'instrument' of the TIMED satellite mission, and the PSMOS, GEM and S-RAMP programs.
- Present the results of our scientific work at international fora.
- Continued presence on the SCAR STAR committee.
- Participate in the activities of NDSC (Network for Detecting Stratospheric Change).
- Establish a web site with on-line synoptic atmospheric, ionospheric and magnetospheric data.

- Provide quality calibrated data to the World Data Centre, IPS Radio and Space Services and others on a timely basis (for space weather and other studies).

31.3 Goal 2 – To protect the Antarctic environment

The Madrid Protocol identifies specific types of damage to the Antarctic environment that should be avoided:

- adverse effects on climate or weather patterns;
- significant adverse effects on air or water quality;
- significant changes in the atmospheric, terrestrial, glacial or marine environments;
- detrimental changes in the distribution, abundance or productivity of species or populations of species of fauna and flora; and
- further jeopardy to endangered or threatened species; or degradation of, or substantial risk to, areas of biological, scientific, historic, aesthetic or wilderness significance.

Assessing these involves monitoring changes in the environment against a background of natural variability. Understanding and quantifying changes can be enhanced significantly by studying the middle and upper atmosphere. Human activity has already caused significant stratospheric ozone depletion and may have contributed to global tropospheric warming. Reported cooling in the mesosphere (near 87 km), where more noctilucent clouds (NLC's) are now observed, may be related to oxidation of methane (a greenhouse gas) at these altitudes.

Research in this output group must satisfy two criteria – (1) it must investigate processes that are directly related to human activity (e.g. greenhouse gas emissions; production of ozone depleting chemicals; production of gaseous pollutants from stations, ships, aircraft), and (2) it must be used in international fora to control these activities. The human activity of relevance to our research is the emission of greenhouse gases and work progresses toward determining if this is having an effect on the climate of the middle and upper atmosphere. At present, measurements of anthropogenic gas concentrations are not made at any of the Australian Antarctic stations. This shortcoming should be redressed to allow the determination of the importance of minor atmospheric constituents on climate: water vapour is the most significant greenhouse 'gas'; aerosols play a major role in ozone depletion chemistry; and volcanic aerosols play a role in cooling the troposphere.

31.3.1 Key scientific output 2.1 – Measurements of atmospheric trace gas concentrations and effects

Undertake measurements of atmospheric trace gas concentrations, specifically ozone and NO_x, using ground-based instrumentation at Davis. This will be part of Australia's response to anthropogenic ozone

depletion in Antarctica, and will contribute to the charter activities of the Network for Detecting Stratospheric Change (NDSC). Investigate the occurrence and physics of NLC's at high southern latitudes, to monitor the effects of atmospheric pollution in the mesosphere. Initiate monitoring studies of Polar Stratospheric Clouds (e.g. with LIDAR) as an adjunct to Madrid Protocol involvement.

31.3.2 Assessment milestones 2.1

- Monitor ozone levels above Davis for one or more solar cycles.
- Investigate the occurrence rate of NLC's at high southern latitudes and determine the relative roles of cooling trends and increased atmospheric emissions.
- Monitor the polar atmospheric environment (stratosphere, mesosphere and thermosphere) to determine whether any adverse effects on climate and weather are observed that might relate to atmospheric emissions.

31.4 Goal 3 – Understand the role of Antarctica in the global climate system

The association of tropospheric climate and solar activity in the pre-industrial era leads to the inference that approximately one third of the global warming observed through the twentieth century results from increased solar activity. The geoelectric field is postulated to provide a mechanism via which solar variability can influence weather and climate. Antarctica provides a necessary platform for testing this hypothesis.

The measurement of middle atmospheric parameters has only recently been achieved, so that the investigation cycle of measurement, characterisation and interpretation is at an early stage. Our measurements are guided by what is already known but much remains to be learned from these measurements. The MFSA radar, the OH spectrometer and the Fabry-Perot spectrometers that are operated in the Australian Antarctic Territory can measure winds through the height range 60–100 km and temperatures near 86 km. The LIDAR system will measure winds and temperatures through the stratosphere and lower mesosphere. Weather balloons provide wind and temperature data in the troposphere and lower stratosphere. Hopefully, the balloon data set limitations, caused by the small number of flights launched per day (one or two) will be overcome by installing a VHF radar that can measure wind speeds minute by minute. Satellites can make global measurements of many parameters but are very limited in investigations at time scales less than a month. Thus ground-based instruments form a vital part of satellite programs as shown by the upcoming TIMED (Thermosphere Ionosphere Mesosphere Energetics and Dynamics) program whose payload is due for launch in 2001. TIMED has included the collection of collaborating ground based instruments as the missions 'fifth instrument'.

31.4.1 Key scientific output 3.1 – Middle atmosphere climate change indicators

Enhanced CO_2 and methane levels in the troposphere have the effect of improving the containment of heat near the Earth's surface. They also make the process of radiating the stratosphere's heat into space more efficient. Both processes make the stratosphere and mesosphere cooler. Long-term measurements of the temperature in the stratosphere at low latitudes have shown a cooling trend consistent with this hypothesis. By measuring temperatures in the mesosphere with the OH spectrometer and in the stratosphere with LIDAR, it will be possible to quantify this trend in the polar regions. The intense cooling of the summer polar mesosphere creates an environment where ice clouds (known as noctilucent clouds) can grow, however, the observational data set in the high northern latitudes only extends back through the last century. They are thought to be a manifestation of climate change but a dearth of observations at southern latitudes may show up a hemispheric asymmetry in mesospheric temperatures. The study of noctilucent clouds and the associated radar phenomena of Polar Mesospheric Summer Echoes (PMSE's) are of interest in climate change studies.

The key factors in the development of the ozone hole are the enhanced destruction of ozone facilitated by Polar Stratospheric Clouds (PSC's) and the barrier between the ozone hole and the ozone rich air to the north that is formed by the stratospheric polar vortex. Located at the northern edge, Davis is ideal to investigate the dynamics and springtime decay of the polar vortex. These investigations, in conjunction with measurements of trace gases and aerosols, will contribute to investigations of the role ozone plays in modifying regional and global climate.

31.4.2 Assessment milestones 3.1

- Make long term measurements of the stratospheric temperature over Davis using the LIDAR. From these measurements, quantify trends in polar stratospheric temperatures.

- Continue the collection and analysis of hydroxyl (OH) spectra. From these spectra, and the associated calibrations, produce a data set of temperatures at the height of the OH layer.

- Make high resolution LIDAR measurements of the vertical extent and frequency of occurrence of tropospheric cirrus clouds, PSC's and NLC's, with concurrent wind and temperature measurements.

- Monitor stratospheric aerosol loading and its temporal characteristics using LIDAR.

- Use spectrophotometer measurements at Davis to obtain column abundances of trace gases and investigate their role in stratospheric climate.

- Identify noctilucent cloud events and make concurrent wind and temperature measurements using the MFSA radar, OH spectrometer, and VHF radar.

- Search for signatures of PMSE's in the MFSA radar data. Search methods need to be developed for both the variable and constant gain systems that have operated at Davis.

- Search for signatures of PMSE's in the VHF radar data. Resolve northern/ southern hemisphere differences in PMSE's.

31.4.3 Key scientific output 3.2 – Gravity wave climatology, propagation and dissipation studies

Gravity waves play an important role in transporting momentum throughout the middle atmosphere and can be characterised through studies of the horizontal winds obtained from the Davis MFSA radar. The variance of the wind speed is related to the wave energy and, with the extensive height range that is sampled by the radar, can be characterised at a number of heights. Combined with the wind measurements themselves, it is possible to study the propagation of gravity waves through the atmosphere. Wind measurements obtained using the LIDAR will allow AGW studies to be extended into the stratosphere where few measurements have been made, apart from radiosonde studies of the lower stratosphere. A VHF radar would allow the gravity wave climatology to be extended into the troposphere. This would enable sources to be investigated. The capability of the LIDAR to measure temperatures, in combination with its wind measurements, may allow the measurement of gravity wave heat fluxes. Spectral studies of the horizontal velocities and the vertical wavenumber are possible using radar and LIDAR data.

31.4.4 Assessment milestones 3.2

- Produce a climatology of gravity wave activity above Davis from the MFSA radar and LIDAR data. This should be as extensive in height and time, as the data will allow.

- Investigate propagation effects in gravity waves above Davis. Investigate the modulation of gravity wave activity by tides and other planetary scale waves.

- Investigate gravity wave dissipation through studies of turbulence.

- Investigate the role of gravity waves in the vertical transport of constituents. If enhancements of radar echo power in the winter polar stratosphere are linked to the downward transport of NO from the thermosphere, then signatures of enhanced gravity wave dissipation may be evident in the data.

- Extend the height range of these studies using VHF radar technology.

31.4.5 Key scientific output 3.3 – Climatology of planetary waves

Global scale waves of periods from 8 hours to many days are evident through the troposphere, stratosphere, mesosphere and lower thermosphere. Observations of winds and temperatures in the mesosphere and lower thermosphere, using the MFSA radar, OH spectrometer and Fabry-Perot spectrometer, and in the stratosphere using

the LIDAR, will allow the measurement of planetary waves and tides over Davis. These measurements may also allow the discovery of new modes, as evidenced by the recent identification of a 10 hour planetary wave in data collected at South Pole. The geographical extent of the polar modes and their interaction with lower latitude modes is not understood. Interactions between tides and gravity waves apply a forcing on the tide that can change its character on a short time scale. Investigations of the variability of planetary waves and tides, and the role that gravity waves play in distorting them are possible with the suite of instruments at Davis.

31.4.6 Assessment milestones 3.3

- Produce a climatology of planetary waves and tides in the mesosphere and lower thermosphere above Davis using the radar and optical techniques.

- Produce a climatology of planetary waves and tides in the stratosphere above Davis using the LIDAR system.

- Characterise the 10 hour and other polar modes in the mesosphere data as a function of height and time.

- Investigate tidal variability in MFSA radar data and the role that gravity and planetary waves play in producing that variability.

31.4.7 Key scientific output 3.4 – Determine the viability of solar variability influencing the weather and climate via modulation of the geoelectric field

Thunderstorm activity draws current upward from the Earth. The lower ionosphere (above ~ 80 km) disperses an estimated 50% of this current globally and returns it to the Earth via that portion of the Earth unaffected by thunderstorm activity. Solar activity significantly influences the conductivity of the atmosphere in the crucial region between the top of thunderstorm clouds and the lower reaches of the ionosphere by modulating cosmic ray ionisation. Thunderstorm development is hypothesised to depend on the ease with which current can be discharged to the ionosphere.

Antarctica, because of its high average elevation and its low absolute humidity is strongly coupled to the global geoelectric circuit. Additionally, because the Earth's magnetic field funnels solar charged particles into the polar atmospheres, the atmosphere above Antarctica is significant in solar variability influences on the geoelectric circuit. Understanding Antarctica's influence on the geoelectric circuit is vital to determining the viability of solar variability influencing weather and climate via modulation of the geoelectric field.

31.4.8 Assessment milestones 3.4

- Determine the viability of the hypothesis that solar variability influences weather and climate through modulation of the geoelectric field, as possible, by analysis of a Vostok geoelectric field data set collected from 1998. Determine the viability of the hypothesis that solar

variability influences weather and climate through modulation of the geoelectric field.

- Expand a cooperative, international, geoelectric field observatory network through the ATS to test whether a solar variability influence on weather and climate via modulation of the geoelectric field exists.

31.4.9 Key scientific output 3.5 – Upper atmosphere climate change indicators

Routine observations made in Antarctica form a climatological baseline extending back to 1958 at Mawson and Casey, and for almost two decades at Davis and Macquarie Island. Because the upper atmosphere changes rapidly from one location to another, each of the four sites makes a unique contribution to the climate archive. Data are analysed and the results are stored in World Data Centres. Both the processed and raw data are available to other researchers, and on the Internet.

31.4.10 Assessment milestones 3.5

- Data processed, validated, and stored in World Data Centres.
- Raw data available for world community.
- Improved empirical models of the ionosphere.
- Validation of physical models of the ionosphere.
- Continued development of a high latitude climate archive.

31.5 Goal 4 – To undertake scientific work of practical, economic and national significance

The geographic and magnetic locations of Australia's Antarctic stations provide a significant opportunity to contribute to space weather monitoring and investigations. The extension of space-based technological systems has made space weather research more important. For example, the energy that a single solar disturbance puts into the geospace environment can destroy orbiting satellites, disrupt continental power grid systems, radio communications (including satellite TV and mobile telephones) and GPS navigation. Space weather research involves the development of an operational global circulation model linking the magnetosphere-ionosphere-atmosphere systems. These models, using real time data inputs, are used to minimise the impact of potentially damaging space weather events.

Ionospheric convection will be measured using a DPS-4 ionosonde at Casey and Davis, a CADI ionosonde at Casey and Macquarie Island, and the Tasman International Geospace Environment Radar (TIGER) at Hobart. TIGER will provide ionospheric convection maps over a wide area. The DPS-4 will give height profiles of convection through the thermosphere. At present, the Australian Antarctic Division and its collaborators are world leaders in determining the extent to which convection is different at different thermospheric heights. The dynamics of the thermosphere is studied using FPS's at Mawson and Davis.

We have shown the power of combining observations from several stations and in particular Mawson and Davis, where there is an overlap of the field-of-view. These stations are also important for comparing thermosphere-ionosphere motions with ionospheric convections measured by the Syowa SuperDARN radar. We have also identified rapid upwellings within the polar cap and are seeking a satisfactory explanation of this phenomenon.

31.5.1 Key scientific output 4.1 – Space weather nowcasting

The study of geospace from Antarctica is important for the following listed reasons.

(i) The geomagnetic and geographic poles are much further apart than they are in the Arctic, allowing the separation of phenomena that are under solar and geomagnetic control.

(ii) Geomagnetic field lines, with their footprints on the Antarctic, connect to the outer regions of the magnetosphere and interplanetary space. As lines of communication, they allow us to remotely and inexpensively monitor these regions from the ground.

(iii) The Arctic and Antarctic regions are interconnected by geomagnetic field lines so that our research contributes to conjugate research.

(iv) The environment is pristine and untouched allowing low noise level electromagnetic and optical observations to be undertaken.

(v) The data from radar, geomagnetic and optical instrumentation provides important input for the modelling and simulation of space weather.

The proposed CRC on Satellite Systems will launch FedSat late in the year 2001. With a low Earth polar orbit, the data gathered from the solar wind through the magnetosphere to the ionosphere will integrate with ground-based observations at the Australian Antarctic stations to provide near-real time space weather data.

31.5.2 Assessment milestones 4.1

- The program of basic observatory operation to service international commitments and participation in global multipoint studies will continue.
- Utilisation of Antarctic and Macquarie Island digital ionosondes to provide space weather data.
- Utilisation of geomagnetic and optical data from all Australian Antarctic stations to provide space weather data.
- Operation of the TIGER SuperDARN radar and utilisation of real-time data to contribute to space weather nowcasting.
- Provide real time ionospheric velocity maps from TIGER for space forecasting.
- Conduct international collaborative campaigns with TIGER.

- Utilise SHIRE to study magnetic impulse events and the drift of plasma patches above Davis.

- Undertake simultaneous FedSat and ground magnetometer correlation studies.

- Continue FPS thermospheric 630 nm observations at Mawson and Davis, and extend the database and enhance circulation models.

- Publish papers on climatology of thermospheric winds and temperatures at Mawson and Davis, and use these to test thermospheric models.

- Provision of thermospheric wind and temperature data for validation studies for TIMED and SABER satellite missions (from Davis LIDAR and Czerny-Turner spectrometer, and Mawson and Davis FPS's).

- Develop prototype automated scanning FPS systems to replace aging units at Mawson and Davis to facilitate real time model validation.

31.5.3 Key scientific output 4.2 – High frequency radio communication forecasting

Data from the Antarctic stations can give immediate support for high frequency radio propagation and are therefore of direct benefit to Australian users. These data are used to generate real time ionospheric models, some of which are made available on the Internet in real time. Data from riometers can be used to identify polar-cap absorption events, which can prevent high frequency radio systems from working. Particle effects, monitored by high-latitude riometers, especially at Mawson and Macquarie Island, can be used to monitor the environment in the vicinity of geostationary satellites. These observations will be compared with space based observations to establish the validity of this prospect, which could offer an effective, low cost method for monitoring the environment of the Australian geostationary and orbiting space-based assets (FedSat1).

31.5.4 Assessment milestones 4.2

- High frequency radio support for the Australian Antarctic region based on real time ionospheric observations.

- Confirmations of polar-cap absorption events.

- A potential monitor for Australian geostationary orbit environment.

- Upgrade ionosondes and improve real time data handling.

- Increased use of high-latitude observations in the Australian Space Forecasting Centre.

- Development of improved high-latitude ionospheric models.

31.5.5 Key scientific output 4.3 – Practical value basic science

GPS receivers at the Australian Antarctic stations provide continuous measurements of the spatial and temporal variation in the total electron content of the southern high-latitude ionosphere. The TIGER radar

will provide location data on the southern auroral oval that folds into the JORN surveillance radar beam. Magnetometer observations are necessary for the production of navigation charts when incorporated into Australian developed magnetic field models. The daily, seasonal and sunspot-cycle variability of ionosphere parameters provide invaluable data for improving forecasting models of the polar region. Applications include corrections to altimeters for sea height studies, study of effects mapped down from the solar wind to the ionosphere and the middle atmosphere, and disturbances propagated to lower latitudes from the high-latitude ionosphere (e.g. gravity waves).

31.5.6 Assessment milestones 4.3

- Maintain basic geophysical observations at ANARE stations.
- Utilisation of magnetometer observations for the production of navigation charts.
- Studies of scintillation effects on propagation from communication and navigation satellites.
- Provide basic geophysical data from each of the ANARE stations for use by scientists of all nations, and to provide real-time data where it is of practical importance.
- TIGER auroral oval location is provided to the JORN surveillance project.
- Utilisation of ionospheric data for improving models of the polar region.

31.6 Conclusion

The strategic plan for Australian Antarctic middle and upper atmospheric physics as presented in this paper is currently being implemented. The Australian Antarctic Division, Atmospheric and Space Physics program at Davis station will continue to be the hub for both middle and upper atmosphere climatology, and space weather research. Mawson and Macquarie Island will operate automated physics observatories from 2001, as will Casey from 2002. Ongoing automation and telescience at the Antarctic and sub-Antarctic stations will ensure near real time data acquisition and satellite retrieval to allow analysis, distribution and archive of the data. The implementation of the strategic plan as presented in this paper, will ensure that Australia remains at the forefront of Antarctic middle and upper atmosphere physics research.

31.7 Acknowledgements

We acknowledge the contribution to the strategic plan by the following members of the Australian Solar-Terrestrial and Space Physics community: Dr Gary Burns, Dr Andrew Klekociuk, Dr John Innis, Dr Anthony Breed and Dr Pene Greet (Australian Antarctic Division); Dr Phil Wilkinson (IPS Radio and Space Services); Associate Professor Bob Vincent (University of Adelaide); Professor Peter Dyson and

Dr Elizabeth Essex (La Trobe University); Professor Brian Fraser (Newcastle University) and Dr Charlie Barton and Mr Stewart Dennis (Australian Geological Survey Organsiation).