ELSEVIER

Journal of Atmospheric and Solar-Terrestrial Physics 66 (2004) 381–392

Journal of ATMOSPHERIC AND SOLAR-TERRESTRIAL PHYSICS

www.elsevier.com/locate/jastp

Retrieval of mesospheric electron densities using an optimal estimation inverse method

J. Grant[a,*], R.G. Grainger[b], B.N. Lawrence[c], G.J. Fraser[a], H.A. von Biel[a], D.N. Heuff[a], G.E. Plank[a]

[a] *Department of Physics and Astronomy, University of Canterbury, Private Bag 4800, Christchurch 8002, New Zealand*
[b] *Department of Atmospheric Oceanic and Planetary Physics, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, UK*
[c] *Rutherford Appleton Laboratories, Didcot, UK*

Received 5 February 2003; received in revised form 18 December 2003; accepted 31 December 2003

Abstract

We present a new method to determine mesospheric electron densities from partially reflected medium frequency radar pulses. The technique uses an optimal estimation inverse method and retrieves both an electron density profile and a gradient electron density profile. As well as accounting for the absorption of the two magnetoionic modes formed by ionospheric birefringence of each radar pulse, the forward model of the retrieval parameterises possible Fresnel scatter of each mode by fine electronic structure, phase changes of each mode due to Faraday rotation and the dependence of the amplitudes of the backscattered modes upon pulse width. Validation results indicate that known profiles can be retrieved and that χ^2 tests upon retrieval parameters satisfy validity criteria. Application to measurements shows that retrieved electron density profiles are consistent with accepted ideas about seasonal variability of electron densities and their dependence upon nitric oxide production and transport.

Keywords: Radar; Partial reflections; Mesosphere; Electron density; Optimal estimation retrieval

1. Introduction

The sampling of partially reflected medium frequency (MF) radar pulses, has, since 1953, been the main ground-based method of determining D-region electron densities. Electron density ($N_e(z)$) profiles have been retrieved using either the differential absorption (DA) method (e.g. Gardner and Pawsey, 1953; Belrose and Burke, 1964; Gregory and Manson, 1969; von Biel, 1995, 1999) or the differential phase (DP) method (von Biel et al., 1970; Austin, 1971).

The latter method relies on determining the difference in the amount of Faraday rotation undergone by the two magnetoionic modes (ordinary mode [o-mode] and extraordinary mode [e-mode]) formed when radar pulses interact with the ionosphere. The major limitation of this method is that it is difficult to quantify the large number of rotations of the modes at upper heights of the D-region (Austin, 1971).

Alternatively, the DA method has been much more widely used and has existed in many variants since it was first introduced (Gardner and Pawsey, 1953). A feature common to all DA experiments is the sampling of partially reflected MF pulses in order to determine the ratio of the backscattered amplitudes of the elliptically polarised e and o modes (A_e/A_o). These modes of angular frequency ω are assumed to be absorbed to differing extents as they propagate up to and back down from a reflecting discontinuity, so that comparing A_e/A_o to the modelled ratio of the reflection coefficients of each mode (R_e/R_o) allows the difference in the amount of absorption undergone by each mode (DA) to be determined. The amplitude ratio is

* Corresponding author. Fax: +64-3-3642469.
E-mail addresses: john.grant@canterbury.ac.nz (J. Grant), r.grainger@atm.ox.ac.uk (R.G. Grainger), B.N.Lawrence@rl.ac.uk (B.N. Lawrence).

1364-6826/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jastp.2003.12.006

given by

$$\frac{A_e}{A_o}(z) = \frac{R_e}{R_o}(z) \exp\left[-\frac{2\omega}{c}\int_{z_o}^{z}(\chi_e(z) - \chi_o(z))\,dz\right],$$

where c is the speed of light, z_o and z, respectively, represent the bottom of the ionosphere and the height from which a signal is reflected, and, $\chi_{o,e}$ are the imaginary parts of the complex refractive indices of the o and e modes. DA is dependent upon both electron concentration and the collisional frequency of electrons with neutral molecules (Gardner and Pawsey, 1953). Evaluating the latter quantity from pressure data in order to obtain the normalised with respect to N_e quantity $(\chi_e - \chi_o)_n$ enables $N_e(z)$ to be determined from

$$N_e(z) = \frac{(d/dz)[\ln((R_e/R_o)(z)) - \ln((A_e/A_o)(z))]}{(2\omega/c)(\chi_e(z) - \chi_o(z))_n}. \quad (1)$$

(Belrose and Burke, 1964; Gregory and Manson, 1969; Flood, 1968; von Biel, 1995, 1999).

The numerator of Eq. (1) predicts that $N_e(z)$ is proportional to the gradient of the DA data profile. In practice, the tendency for this profile to exhibit negative gradients at certain heights yields unphysical negative electron densities. This anomaly constitutes a major weakness in the application of the DA method.

Various explanations for the anomaly have been proposed. The possibility that $A_e/A_o(z)$ is derived from signals contaminated by off-vertical reflections has been suggested (e.g. Holt et al., 1961; Belrose and Burke, 1964; Belrose, 1970). Other researchers have focused upon the possibility that the Fresnel reflection model used to obtain the $R_e/R_o(z)$ profile is insufficiently realistic to account for the physics of pulse ionosphere interactions. Flood (1968) proposes using a model in which backscatter of the incident pulse is achieved by many scatterers within the pulse volume (the *volume scatter* model), with absorption of the pulse by electrons within the volume also being taken into account. von Biel (1999) attempts to overcome the problem by mathematically fitting a monotonically increasing function to the DA data profile, showing that the residuals arising from this fit can be interpreted as discrepancies in the $R_e/R_o(z)$ profiles.

Given the turbulent nature of the mesosphere (e.g. Lawrence and Randel, 1996), it is highly probable that small-scale fine structure consisting of large variations of electron density exists at various heights throughout the D-region (e.g. Thrane and Grandal, 1980; Hocking and Röttger, 1983) and that modelling $R_e/R_o(z)$ profiles does not give realistic representations of the true reflection processes. In an attempt to account for this structural variability and obtain more physically realistic parameterisations of the partial reflections, we outline an alternative to the traditional methods of retrieving mesospheric electron densities.

The new method requires the use of an optimal estimation (OE) inverse method (Rodgers, 1990) in which a *forward model* simulates sensing of the ionosphere by our MF polarimeter radar situated at Birdlings flat (43.8°S, 172.7°E), near Christchurch, New Zealand. Both the absorption and reflection of pulses are, respectively, accounted for by retrieving $N_e(z)$ and $[\sum \Delta N_e](z)$ profiles. The former is interpreted as the quantity which provides attenuation of the propagating signals while the latter is the quantity upon which the amplitude of the backscattered signal from any height depends and is a measure of ionospheric discontinuity. The summation sign indicates that the retrieved quantity is the sum of differential electrons within each height interval sampled by the polarimeter.

Given that the number of discontinuities within a sampling interval is likely to vary with height and that the discontinuities are unlikely to be of the same magnitude, we attempt to account for this by retrieving a third parameter which we interpret as a measure of the distribution of discontinuities within a sampling interval. We obtain this parameter by assuming a Gaussian distribution of ΔN_e values within the sampling interval and specify that the standard deviation (σ_m) of the Gaussian distribution be retrieved, with m representing a fixed number of possible scattering levels within each sampling interval.

A major feature of this retrieval is, that not only does it parameterise the absorption and reflection processes, it also combines modelling of both the absorption and Faraday rotation of each mode.

2. Retrieval overview

An explicit description of the underlying mathematical principles and general notation for the optimal estimation (OE) retrieval is given by Rodgers (2000).

In essence, an atmospheric state vector $\mathbf{x}$ comprising the most probable $N_e(z)$, $[\sum \Delta N_e](z)$ and $\sigma_m(z)$ profiles is found. A forward model relates $\mathbf{x}$ to simulated measurements $\mathbf{y}$ by including the polarimeter response function and the physics which affects the propagating pulse on its journey up to and back from a reflecting layer. From an initial estimate of the atmospheric state, the simulated measurement vector and the real measurements are compared. Subsequent refinements to the state are then made using the Levenberg–Marquardt method, an iterated state $\mathbf{x}_{i+1}$ being given by

$$\mathbf{x}_{i+1} = \mathbf{x}_i + (\mathbf{S}_a^{-1} + \mathbf{K}_i^T\mathbf{S}_\varepsilon^{-1}\mathbf{K}_i + \gamma\mathbf{I}_n)^{-1}(\mathbf{K}_i^T\mathbf{S}_\varepsilon^{-1}[\mathbf{y} - \mathbf{F}(\mathbf{x}_i)] - \mathbf{S}_a^{-1}[\mathbf{x}_i - \mathbf{x}_a]), \quad (2)$$

where $\mathbf{x}_a$ is the a priori state; $\mathbf{S}_a$ and $\mathbf{S}_\varepsilon$ are, respectively, the a priori and measurement covariance matrices; $\mathbf{K_i}$ is the Jacobian matrix of the previous converged state $\mathbf{x}_i$ and $\mathbf{F}(\mathbf{x_i})$ represents the forward model. γ is a variable factor which balances the minimisation between Gauss–Newton and steepest descents convergence (e.g. Press et al., 1992) and $\mathbf{I}_n$ is an $n \times n$ identity matrix, where n is the number of elements in $\mathbf{x}$. Updating of $\mathbf{x}$ ceases when running the forward model no longer effects a significant minimisation

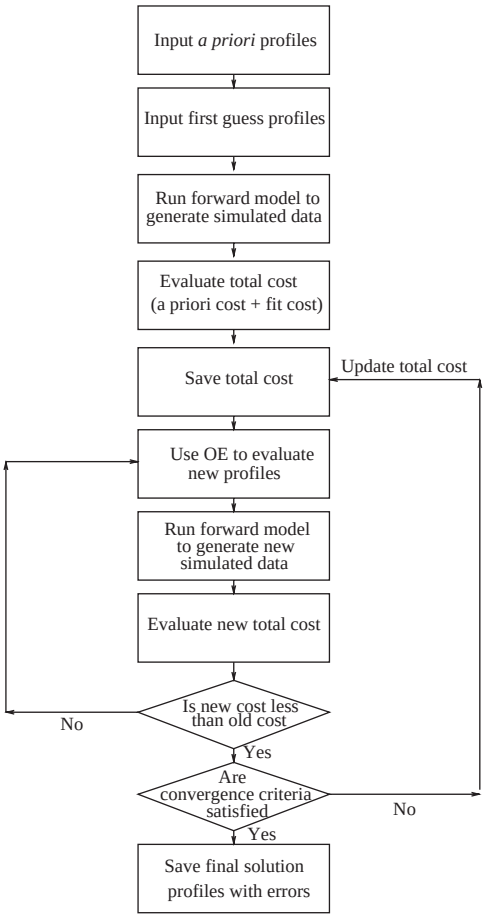

Fig. 1. Flow diagram showing the retrieval process.

of a cost function which is evaluated using

$$\mathbf{g}(\mathbf{x}) = (\mathbf{y} - \mathbf{K}\mathbf{x})^{\mathrm{T}}\mathbf{S}_{\varepsilon}^{-1}(\mathbf{y} - \mathbf{K}\mathbf{x}) + (\mathbf{x} - \mathbf{x}_{\mathbf{a}})^{\mathrm{T}}\mathbf{S}_{\mathrm{a}}^{-1}(\mathbf{x} - \mathbf{x}_{\mathbf{a}}). \quad (3)$$

The overall retrieval process can be regarded as continually making virtual measurements of a state space which is altered until one of those measurements produces simulated data which closely approximates the real data. This process is summarised in the flow diagram of Fig. 1.

3. The forward model

3.1. Propagating wave physics

For a wave of angular frequency ω, which has been partially reflected from an ionospheric discontinuity at a height z within an ionosphere whose lower boundary is at a height z_{o}, the amount of absorption and Faraday rotation undergone by the partially reflected o-mode and e-mode can be quantified in terms of their amplitudes A_{o} and A_{e}, and, their phases ϕ_{o} and ϕ_{e} according to

$$A_{\mathrm{o,e}}(z) = \frac{1}{2z}\,|R_{\mathrm{o,e}}(z)|\exp\left(-\frac{2\omega}{c}\int_{z_{\mathrm{o}}}^{z}\chi_{\mathrm{o,e}}(z)\,\mathrm{d}z\right) \quad (4)$$

and

$$\phi_{\mathrm{o,e}}(z) = -\frac{2\omega}{c}\int_{z_{\mathrm{o}}}^{z}\mu_{\mathrm{o,e}}(z)\,\mathrm{d}z + \beta_{\mathrm{o,e}}(z), \quad (5)$$

$|R_{\mathrm{o}}|$ and $|R_{\mathrm{e}}|$, respectively, represent the magnitude of the o-mode and e-mode reflection coefficients and β_{o} and β_{e}, respectively, represent the reflection phase changes. The quantities μ and χ are, respectively, the real and imaginary parts of the complex refractive index and are evaluated for each mode by applying the Sen–Wyller generalisation (Sen and Wyller, 1960) of the Appleton–Hartree equation.

3.2. Fresnel scattering

The Fresnel backscatter model is implemented by assuming that there will be a number of Fresnel scatterers within a height bin. Despite the fact that we subdivide each 1 km height interval into a fixed number (m) of possible equidistant scattering levels, we can obtain an estimate of the number and magnitude of the actual scatterers by retrieving both the "width" and $\sum \Delta N_{\mathrm{e}}$ of an assumed Gaussian distribution of ΔN_{e} values within the height interval. Our main justification for using a Gaussian is that mathematically, it offers us the simplest way of modelling the distribution of ΔN_{e} whilst at the same time, providing us with a means of varying the number of scatterers by retrieving only the single parameter σ_m from each retrieval height interval. For example, retrieval of a narrow Gaussian (the solid line of Fig. 2) is indicative of there being relatively few scatterers, with ΔN_{e} values at the edge of the Gaussian being small, whilst a broad Gaussian (the dashed line of Fig. 2) indicates that more scatterers are present. To achieve this objective, we retrieve both the standard deviation of the Gaussian (σ_m) and the sum of the differential electrons ($\sum \Delta N_{\mathrm{e}}$) within the sampling interval.

For any sampling layer z, the electron concentration of each of the possible m scattering layers is given by $N_{\mathrm{e}}(z-1) + \Delta N_{\mathrm{e}(m)}$ (see Fig. 2), where $N_{\mathrm{e}}(z-1)$ is the electron density of the sampling layer $z-1$. Application of the Sen–Wyller generalisation to these values enables the complex refractive indices of each mode $n_{\mathrm{o,e}}$ to be obtained for each scattering layer. The complex reflection coefficient profiles $R_{\mathrm{o,e}}(h)$ are now evaluated using

$$R_{\mathrm{o,e}_{(h)}} = \frac{1}{n_{\mathrm{o,e}_{(h)}} + n_{\mathrm{o,e}_{(h-1)}}}\,\frac{\mathrm{d}n_{\mathrm{o,e}}}{\mathrm{d}h}, \quad (6)$$

where $\mathrm{d}n_{\mathrm{o,e}} = n_{\mathrm{o,e}_{(h)}} - n_{\mathrm{o,e}_{(h-1)}}$ and $\mathrm{d}h$ is the distance between scattering discontinuities. This expression is similar to that used by Hocking et al. (1991), who use the approximation of $n_{\mathrm{o,e}} \approx 1$. The subscripts $h-1$ and h refer respectively to a reflecting level and the next reflecting level above it (Budden, 1961).

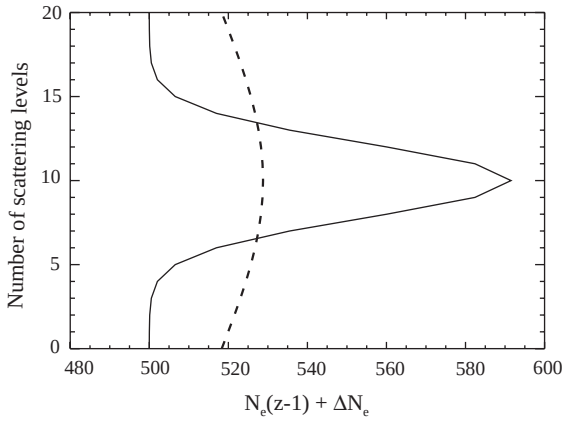

Fig. 2. Fresnel scatterers within a sampling interval. The electron density used to determine the refractive index within each scattering layer is determined by adding $N_e(z-1)$ (the electron density of the sampling interval immediately below. In this case 500 cm^{-3}) to the ΔN_e values. Both the solid and the dashed lines represent possible Gaussian distributions for which $\sum \Delta N_e = 500$.

3.3. Pulse width and power

The argument so far is one which holds to be true in the absence of knowledge about the backscattered power. To ascertain precisely what proportion of the power of the pulse is backscattered to the receivers, it is necessary to simulate the characteristics of the transmitted pulse.

Our polarimeter radar obtains data from 1 km height intervals. However, pulse width effects ensure that signals returning from a sampling level will be the result of superposed signals from many height levels within a pulse width (e.g. Hocking and Röttger, 1983). The possible presence of many Fresnel scatterers within a sampling interval makes it necessary to account for both the transmitted power and the pulse width of our transmitter. The radar equation can be written (e.g. Hocking, 1987)

$$P_r = \frac{P_t G_t}{4\pi(2z)^2} A_\gamma |R|^2, \tag{7}$$

P_t and P_r are, respectively, the transmitted and received powers of the antennae, G_t the gain of the transmitting antenna, A_γ the effective aperture of the receiving and transmitting system, z is the range and R the reflection coefficient. The backscattered power can therefore be evaluated for both the e-mode and the o-mode.

If it is assumed that the transmitter pulse has a Gaussian envelope and a pulse width D at full width of half power, and, the transmitter power is taken into account, the backscattered amplitudes can be evaluated by mathematically convolving the pulse expression

$$g(z) = \sqrt{\frac{P_t G_t A_\gamma}{4\pi(2z)^2}} \exp\left[\frac{(z-z_o)^2}{D^2}\right] \exp[2\pi j\omega(z-z_o)], \tag{8}$$

respectively, with the expressions

$$R'_{o,e}(h) = R_{o,e}(h)\exp\left(-\frac{2\omega}{c}\int_{z_o}^{z} \chi_{o,e}\, dz\right), \tag{9}$$

where $R'_o(h)$ and $R'_e(h)$ represent o-mode and e-mode reflection coefficient profiles which have been adjusted to compensate for the absorption of each mode of the signal and $g(z)$ is adjusted to $g(h)$ where h is a height grid appropriate to the Fresnel scatter model. Since $R'_o(h)$, $R'_e(h)$ and $g(h)$ are complex quantities, the complex backscattered amplitudes are given by (Hocking and Röttger, 1997)

$$A_{o,e}(h) = R'_{o,e}(h) \otimes g(h),$$

where $\otimes$ represents the convolution operation and the range term $1/2z$ is implicit in the amplitude term of the pulse expression.

Eqs. (4) and (5) predict that the backscattered signals are dependent upon the magnitude and phase of reflection coefficients which are the result of a Fresnel reflection mechanism being assumed to be acting at each 1 km height level. However these signals can be interpreted as the resolved magnitude and phase components which result from Fresnel scatter from small scale fine structure present at numerous heights within each 1 km interval. Consequently, we sum the complex amplitude contributions made by the scatterers within the sampling bin to obtain the overall amplitude and reflection phase change occurring within 1 km height intervals.

3.4. Evaluating collisional frequency

All refractive index evaluations are processed for a collisional frequency profile which is given by

$$\nu(z) = A(z) + B(z)\cos\left[\frac{2\pi(\text{Month} - \phi(z))}{12}\right],$$

where z is the height and $\phi(z)$ is the phase (in $\pm$ months from 1 January) at which $\nu(z)$ attains its maximum value. The coefficients $A(z)$, $B(z)$ and $\phi(z)$ are obtained by fitting a fifth-order polynomial to CIRA 86 pressure data (Fleming et al., 1990) for the latitude of Birdlings Flat, with $\nu(z)$ assumed to have a pressure dependence of 68 MHz h Pa^{-1} (von Biel, 1995).

3.5. The radar hardware

The results presented in this paper are from a phase shifted correlation polarimeter (von Biel, 1977). In the current form of the experiment, the carrier frequency of transmitted pulses is 2.4 MHz and 120 bursts of transmitted pulses are produced during a 12 min sampling period each hour. During each burst, 20 pulses with a repetition period of 100 ms are transmitted so that each burst lasts for 2 s with the pulse width being set at 20 μs. For each burst, returning pulses are electronically gated to ensure that ten pulses are received from each of 82 height levels, spaced at 1 km intervals over

a range of 39 to 120 km. The transmitter array consists of four rows, each with two collinear dipoles. The rows are 84 m apart and the input signals to each dipole are in phase so that the transmitted signals are linearly polarised and propagate vertically upward.

The original Gardner and Pawsey experiment (Gardner and Pawsey, 1953) samples the received elliptically polarised modes by inducing opposite senses of circular polarisation in two orthogonal antennae. von Biel (1977) increases the accuracy of the sampling process by continually applying $\pi/4$ phase shift increments to one of two orthogonal antennae. This induces different degrees of elliptical polarisation in the antennae and facilitates more accurate sampling of the total elliptically polarised field since eight measurements of that field are now obtained. In addition to these measurements, switching off the x-antenna and then the y-antenna allows respective measurements of the y-antenna voltage and the x-antenna voltage to be made. The net result is that for each pulse we obtain a single measurement for each induced phase from our 82 height levels. Given that we transmit 120 bursts in 12 min, there are 6 s between equivalent phase samples at any given height.

The data from each height can therefore be used to ascertain all the Stokes parameters associated with the elliptically polarised field obtained from that height, enabling the degree of elliptical polarisation of the field to be determined.

4. A priori and initial estimates

It is necessary to choose a priori $N_e(z)$, $[\sum \Delta N_e](z)$ and $\sigma_m(z)$ profiles which reflect realistic states of the ionosphere. In the case of $N_e(z)$ values, results of previous researchers can be averaged to obtain a profile and variances which can reasonably be regarded as a viable a priori. For $[\sum \Delta N_e](z)$ and $\sigma_m(z)$ profiles, no previous measurements are available and for the former, the choice is restricted to speculating that use of the electron density values will provide a reasonably viable alternative to values based upon actual retrieved profiles. For the latter, we assume a fairly broad Gaussian with σ_m being taken as $m/2$. In this case, it is found that use of the a priori profiles as initial guess profiles readily provides convergence towards a valid solution. The form of the a priori ionisation profiles is

$$N_e(z) = \left[\sum \Delta N_e\right](z) = 50 + 5000 \exp\left[\frac{z - z_o}{H}\right],$$

where z_o is taken to be 90 km, H is taken to be 4 km and the concentration is in electrons/cm^3. It should be noted that these profiles were not altered in any way to account for known seasonal variability, but were used as a priori for all retrievals, regardless of season.

The optimum resolution of the retrieval is found by evaluating *averaging kernels* (see Figs. 3 and 4) using $\mathbf{A} = \mathbf{G_y K_x}$ where the Jacobian $\mathbf{K_x}$ is evaluated after the final iteration and $\mathbf{G_y}$ is the gain matrix, which is a measure of the sensitivity of the retrieval to measurement error. Tests using a variety of grids for the state vector suggest that the optimum resolution of the retrieval is 3 km. Physically, this is likely to be a manifestation of pulse width effects. Consequently, the $N_e(z)$, $[\sum \Delta N_e](z)$ and $\sigma_m(z)$ profiles are retrieved at 3 km intervals.

Accounting for the pulse width of the transmitted signal requires an overlap between the height range of the retrieved profiles and the measurement vectors. Contributions to a measurement signal which is nominally at a certain height can come from regions of the ionosphere both above and

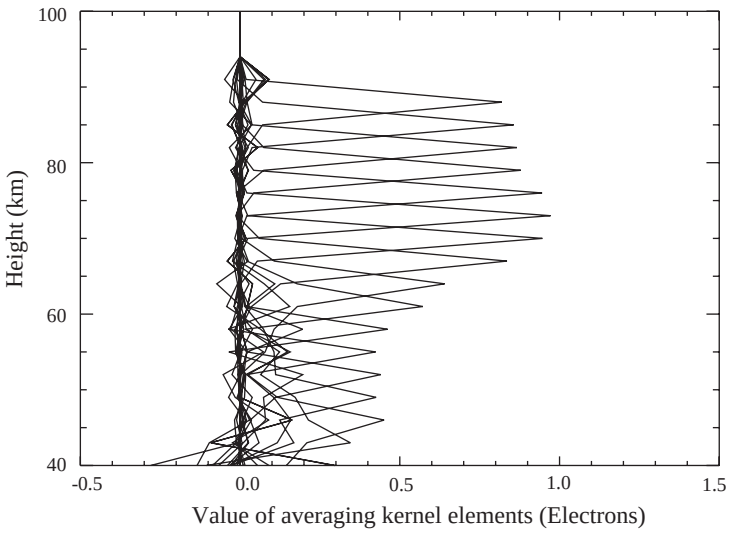

Fig. 3. The averaging kernels for retrieval of a known IRI $N_e(z)$ profile.

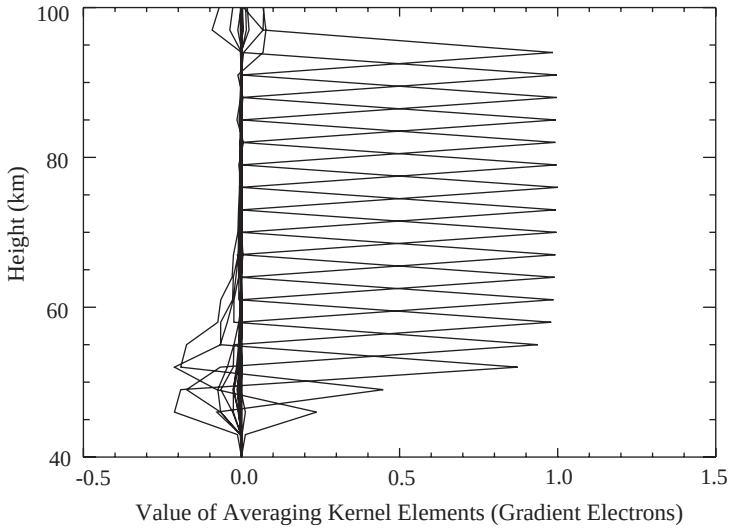

Fig. 4. The averaging kernels for retrieval of a known $[\sum \Delta N_e](z)$ profile.

below that height. We therefore restrict our measurement vector to having a range of 50–90 km. Our experience at Birdlings Flat is that data from below 90 km is not contaminated by sporadic-E and is certainly well below the height at which total reflection from the E-region occurs.

5. Atmospheric noise

The received signal is made up of both elliptically polarised and randomly polarised elements. To isolate the elliptically polarised field which occurs as the result of interactions between a pulse and the ionosphere, it is necessary to evaluate the degree of elliptical polarisation of the field. Whilst it is possible that the randomly polarised field is due to a noise source, an examination of data shows that elliptically polarised noise is also present in the signal.

The Stokes parameters of the received elliptically polarised fields can be used to pre-determine atmospheric noise which is subsequently input to the forward model where it is added to the simulated measurements.

6. Validation of the retrieval

6.1. Retrieval of known profiles

A validation criterion is that the retrieval should be able to reproduce known profiles (see Figs. 5 and 6). To test this, we implement the retrieval using data which has been derived by passing known profiles through the forward model and estimating measurement uncertainties which closely simulate those of real measurements. A NASA International Reference Ionosphere (IRI) (Bilitza, 2001) noon $N_e(z)$ profile for 15 January 1997 is adapted for this purpose. The latest version of the D-region IRI model incorporates the FIRI semiempirical model based upon rocket flight data (Fiedrich and Torkar, 2001) and attempts to account for stratospheric warming effects by including the D-region model of Danilov et al. (1995). The lower boundary of the IRI profile is 67 km and electron densities must be downward extrapolated to 40 km in order for the profile to be used in our retrieval. Our $[\sum \Delta N_e](z)$ values are obtained using the same profile and our $\sigma_m(z)$ values are restricted to $m/2$ by setting their a priori variances to extremely small values.

The averaging kernels for the electron profiles are shown in Figs. 3 and 4. These results are very typical for all retrievals, whether they be of known profiles or unknown profiles. Averaging kernels peaking at values of unity indicate that the solution has been retrieved entirely from the measurements. Reduced peaks indicate that the profiles are partially retrieved from the a priori profiles with smaller peaks indicating the undesirable occurrence of more dependence of the retrieval upon the a priori. The indications are that typically, the retrieved $[\sum \Delta N_e](z)$ profiles are valid for the range of approximately 55–95 km whereas retrieved $N_e(z)$ profiles are valid between approximately 64 and 90 km.

6.2. χ^2 quality control tests

Quality control tests on retrievals which use actual polarimeter data are based on χ^2 statistical tests applied to ensembles of retrieval parameters.

6.2.1. Retrieved measurements against real measurements

To test if the retrieval has converged towards the correct minimum and provided a meaningful solution, the χ^2 formed by the retrieved measurements ($\hat{\mathbf{y}}$) and actual measurement ($\mathbf{y}$) is analysed for yearly ensembles of retrievals. The covariance matrix for this test is formed by an expres-

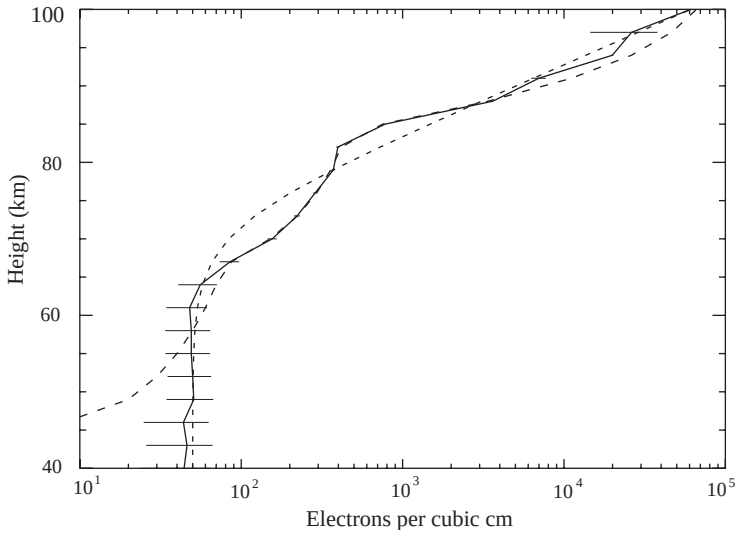

Fig. 5. Retrieved simulated $N_e(z)$ profile (solid line), the known profile (dashed line) and the a priori profile (dotted line).

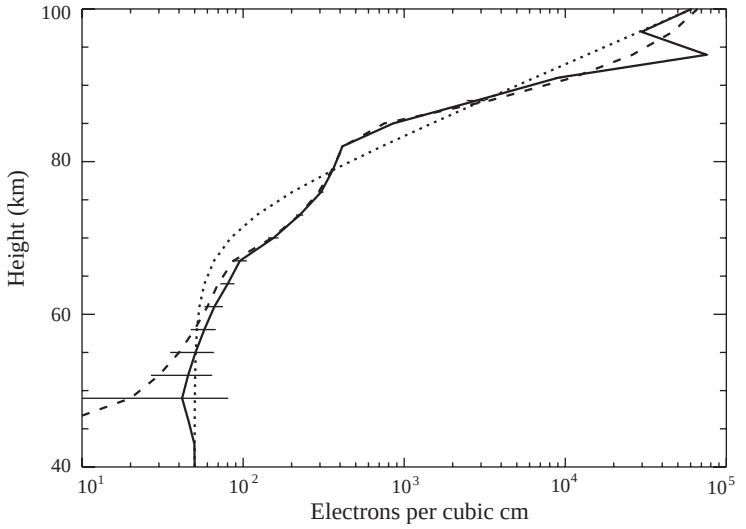

Fig. 6. Retrieved simulated $[\sum \Delta N_e](z)$ profile (solid line), the known profile (dashed line) and the a priori profile (dotted line).

sion which depends upon both the covariance matrix for the measurements and the a priori covariance matrix (Rodgers, 2000, Eq. (12.9)). The number of degrees of freedom for this test is $m-n$, where m represents the number of elements in the measurement vector and n represents the number of elements in the state vector. In this case, $m=320$ and $n=63$. It is desirable therefore that the mean of all the χ^2 values ($\overline{\chi^2}$) is 257.

The results of this validation test show that in all years, $\overline{\chi^2}$, although not being equal to 257 (see Table 1), are reasonably acceptable. The limits of acceptance for χ^2 values are taken to be that they should not exceed twice the yearly $\overline{\chi^2}$ value for yearly ensembles of retrievals.

6.2.2. *Real measurements against a priori measurements*

To carry out this test, the a priori state vector is run through the forward model to produce an a priori measurement vector ($\mathbf{y}_a$). The χ^2 function is formed by $\mathbf{y}-\mathbf{y}_a$ and the test is done for m degrees of freedom so that a desirable $\overline{\chi^2}$ value for this test is 320.

Table 1
Mean yearly χ^2 values for the years 1994–1999

Year	1994	1995	1996	1997	1998	1999
$\overline{\chi^2}$ for $\mathbf{y}-\hat{\mathbf{y}}$	246.7	213.0	261.1	239.7	318.5	291.3
$\overline{\chi^2}$ for $\mathbf{y}-\mathbf{y_a}$	280.9	276.0	334.0	352.2	297.6	302.8
$\overline{\chi^2}$ for $\hat{\mathbf{y}}-\mathbf{y_a}$	25.7	29.6	29.8	21.4	21.5	25.3

Table 2
The number and percentages of retrievals satisfying all the validation tests for the years 1994–1999

Year	1994	1995	1996	1997	1998	1999
Total retrievals	1265	1205	1145	1203	1019	812
Valid retrievals	1126	1100	1014	1122	917	701
Percentage valid	89.0	91.3	88.6	93.3	89.9	86.3

6.2.3. *Retrieved measurements against a priori measurements*

In this case, the retrieved measurement vector $\hat{\mathbf{y}}$ is compared to the a priori measurement vector $\mathbf{y}_a$. This test can only be carried out on measurements from which the retrieved solution is derived. To ensure this, a singular value decomposition is applied to a transformation of the final Jacobian so that a special form of the χ^2 for $\hat{\mathbf{y}}-\mathbf{y}_a$ is evaluated (Rodgers, 2000, c.f. problem 12.1).

The $\overline{\chi^2}$ value for this test is expected to be approximately the same magnitude as the number of state vector elements retrieved from the measurements (approximately 32).

6.3. χ^2 *tests summary*

The number of retrievals which satisfy all the validation tests are outlined in Table 2. The total retrievals in the first line are the total number of retrievals which satisfy two criteria. Firstly, they satisfy polarised noise acceptance criteria we base upon a five year climatology of evaluated polarised noise power we have compiled, and, secondly, they successfully converge to a solution. The indication is that approximately 90% of retrievals are successful each year.

6.4. *Validation against DA*

In addition to the validation tests outlined above, it is useful to assess our retrieval method against the DA method. For this purpose, we rearrange Eq. (1) and invert the DA integral,

$$\frac{2\omega}{c}\int_{z_o}^{z}(\chi_e(z)-\chi_o(z))_n N_e(z)\,dz = \ln\frac{R_e}{R_o}(z) - \ln\frac{A_e}{A_o}(z) \quad (10)$$

iterating to a solution by continually updating the $N_e(z)$ profile in a manner similar to our OE method. We avoid the problem of obtaining negative electron densities by solving the problem in $\log_{10}$ space.

The results of this inversion are shown as the dot-dash lines of Figs. 7 and 8. These represent the monthly mean noon DA electron densities for January 1997 and June 1997, respectively, and show some characteristics which are similar to our OE retrieved profiles.

A general lack of winter OE electrons below 70 km is largely reproduced in the winter DA profile. However, high altitude winter DA values are very much smaller than the winter OE values and do not exhibit the 82 km ledge and the high 75–82 km values seen in the OE profiles.

Summer DA values are generally in better agreement with the OE values. Between 61 and 75 km, the profile shapes are similar. The DA values however, indicate a prominent ledge at 76 km which is not present in the OE profile and do not reproduce the 82–85 km ledge seen in the OE profile. The narrower height range of the DA summer profile is a consequence of the large low altitude electron densities absorbing the e-mode signal. Since the DA method relies in quantifying the difference in absorption of the two modes, it's application requires an upper limit at the height where the e-mode has been completely absorbed.

7. OE retrieval results

7.1. *The data set*

A database of retrievals is compiled from polarimeter data for the years of 1994–1999 inclusive. For this period, a daily data set comprising hourly measurements for 10:00–14:00 inclusive is used to obtain hourly profiles. The retrieved profiles are used to obtain daily mean noon profiles which are further used to determine monthly mean noon profiles. In advance of presenting a climatology of retrieved electron densities in a later paper, we initially present monthly means of retrieved profiles for January and June of 1997.

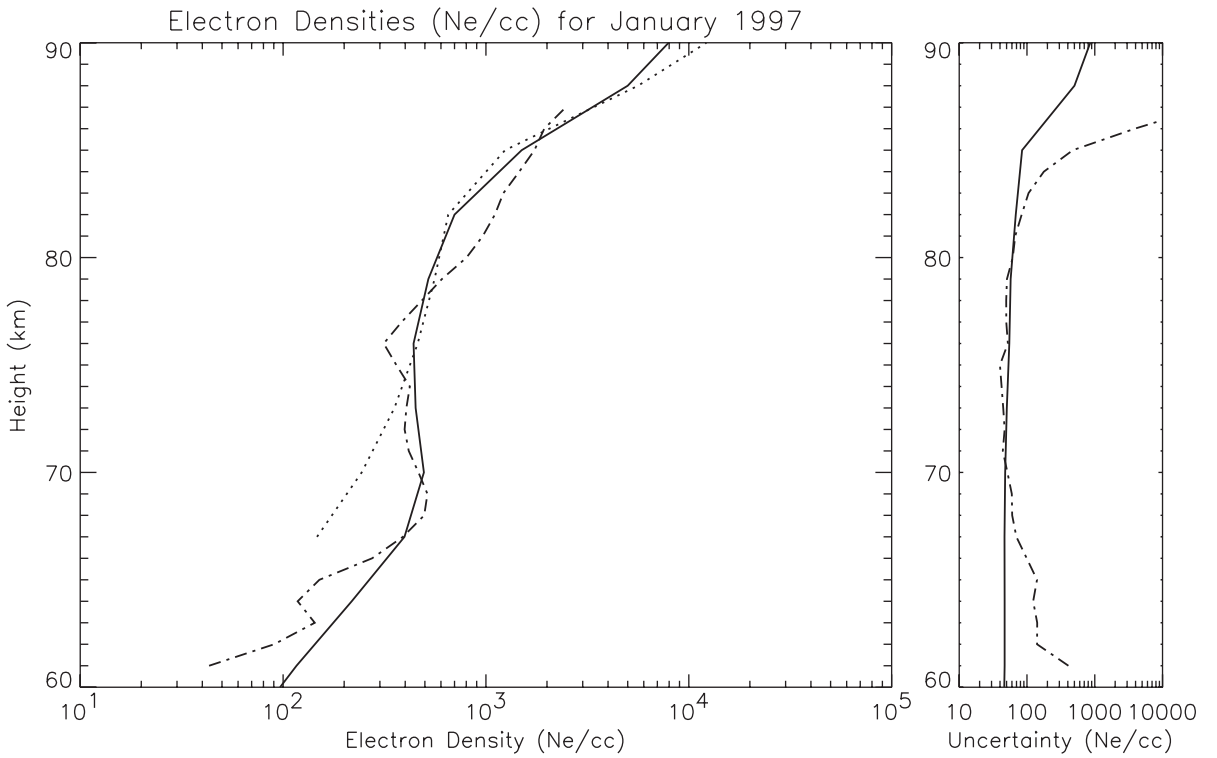

Fig. 7. Monthly mean OE (solid line) and DA (dot-dash line) retrieved noon electron densities for January (summer) 1997 compared to 15 January 1997 IRI noon electron densities (dotted line). The plot on the right displays the monthly mean retrieved uncertainty.

7.2. *Monthly mean profiles*

The monthly mean profile of retrieved noon electron densities for January 1997 (southern hemisphere summer) is shown in Fig. 7 (solid line). This is compared to the mid-January 1997 IRI electron density profile. The major feature of the IRI profile is the ledge between 82 and 85 km. Above 85 km, electron densities increase rapidly with height. This increase is largely reproduced in our OE retrieved profile and has been observed in previous studies of both rocket borne measurements (e.g. Mechtly et al., 1972; Hocking and Vincent, 1982) and DA radar measurements (e.g. Belrose and Burke, 1964; Flood, 1968; Belrose, 1970).

Between 75 and 85 km, there is very good agreement between the profiles, but between 85 and 95 km, the summer IRI values are much larger than the retrieved values. Below approximately 75 km, the retrieved values exceed the IRI values by a significant amount. Nevertheless, the general shape of the retrieved profile is in good agreement with that of the IRI profile.

The equivalent profiles for June 1997 (southern hemisphere winter) are shown in Fig. 8. The general shape of the retrieved profile (solid line) is, in this case, less in agreement with the IRI profiles. Below approximately 72 km, the retrieved values are smaller than the IRI values. Between 72 and 88 km, our retrieved values are much larger than the IRI values and display a convex shape between 72 and 88 km. It is still notable, however, that between 82 and 85 km, there is again evidence of a ledge, above which, there is a large gradient in electron densities.

To assist in a comparison between summer and winter, the mean January (summer) profile (dashed line) is included in Fig. 8. Below 75 km, the January values are appreciably larger than the June (winter) values, in line with solar-zenith angle control. However, between 75 and 82 km, June values exceed January values, with there being reasonably good agreement between the profiles above 82 km. The indication is that this upper region of the profile is not so strongly dependent upon radiation intensity but is more controlled by dynamical influences, these taking the form of enhanced wintertime downward transport of aurorally produced thermospheric NO (e.g. Gregory and Manson, 1969; Garcia et al., 1987; Siskind et al., 1997) or planetary wave advection of NO from the polar night region (e.g. Thomas, 1961; Schwentek, 1974; Kawahira, 1984). In the southern hemisphere winter, it has been observed that increases in ionisation maximise between 80 and 83 km (Ferguson, 1981).

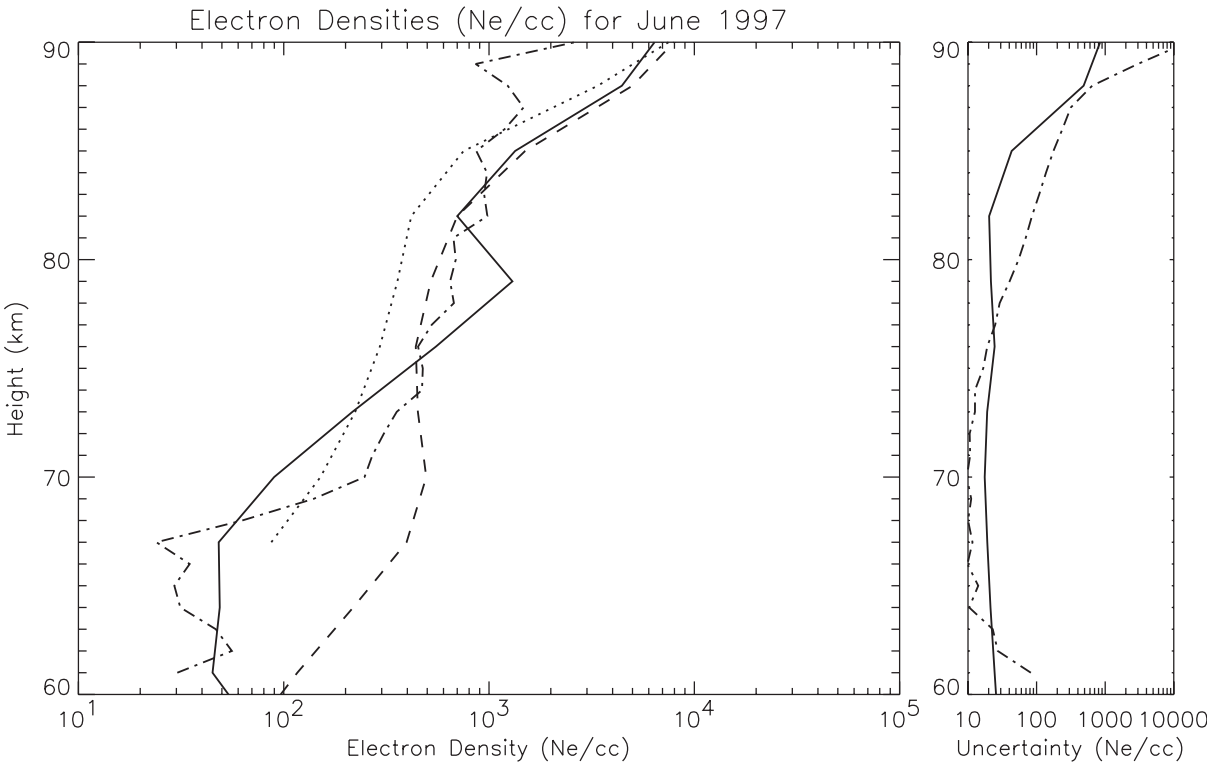

Fig. 8. Monthly mean OE (solid line) and DA (dot-dash line) retrieved noon electron densities for June (winter) 1997 compared to 15 June 1997 IRI noon electron densities (dotted line). To assist comparison, the mean January noon electron densities are included (dashed line). The plot on the right displays the monthly mean retrieved uncertainty.

8. Discussion and conclusions

Although validation of our OE retrieved profiles against DA profiles indicates that there are certain similarities, there are also significant differences. The DA method relies upon evaluating the ratio of the e-mode and o-mode reflection coefficients at the 1 km measurement resolution of the polarimeter radar. It is highly likely that this fails to account for the probable small scale structure which arises from the great deal of turbulence found in this region of the atmosphere. Although our variant of the DA method overcomes the problem of obtaining negative electron densities, it does so at the cost of a poor fit to the DA data profiles where DA data values decrease with height.

The use of a Fresnel scatter mechanism within our forward model must be commented upon. It has been shown that both Fresnel and isotropic scatterers may coexist at many atmospheric heights but that above 80–85 km, kinematic viscosity is likely to be so large that the dominant scattering mechanism is almost certainly more isotropic and associated with turbulence (e.g. Hocking and Vincent, 1982; Hocking, 1987). It is likely that at these heights, our Fresnel scatter model produces simulated signals which are in error. To overcome this problem and build a turbulent reflection mechanism into our forward model, it would be necessary to find the heights of dominant turbulent reflection by using an aspect sensitive radar.

The increase with height of kinematic viscosity also leads to Fresnel scatter which is less specular in nature where the step depth is other than zero (e.g. Hocking, 1987). This indicates a more gradual change in the refractive index at the reflecting boundary. The backscattered power from such a boundary has been investigated by Hocking and Röttger (1983), who show that as the step depth increases, the amount of reflected power is reduced significantly. To model the non-zero step depth dependency would require that the distance between the possible scattering levels within a sampling interval be reduced significantly such that this distance is a fraction of the step depth. An assumed profile describing the variation of N_e over the step would then be required, with the step depth being allowed to vary.

Such an approach would require that a prohibitively large number of possible scattering levels be introduced and that the step depth of the reflecting level be an additional retrievable parameter. The former of these options would dramati-

cally restrict the speed of the retrieval process and the latter, in addition to slowing the retrieval still further, is likely to render the retrieval problem underdetermined.

The position of scatterers within a sampling layer must also be considered. Our forward model, in assuming a Gaussian of ΔN_e values, centres the largest value (the peak of the Gaussian) at the middle of the sampling layer. In reality, the largest scatterer can be anywhere within the sampling layer. Since the phase of the backscattered signal returning to the polarimeter is critically dependent upon the position of the scatterers (e.g Hocking and Röttger, 1983), the reflection phase change returned by our forward model is a possible source of error.

The accuracy of retrieved ΔN_e values depends upon accurately determining the radar power. Altering the power within the forward model has been found to have little or no effect on retrieved N_e values. To be completely sure that the radar power is accurately parameterised within the forward model, it will be necessary to calibrate the radar completely. This is not a trivial task and requires that the transmitted power be accurately measured in order to account for variability of ground reflectivity. The most likely source of such variability would be winter snow. Fortunately, the Birdlings Flat site is not prone to such events even in the most severe winter.

In our present situation, it is technically unfeasible to carry out such a calibration. However, methods exist whereby it would be possible to compare the strength of successive E-region reflections to ground reflectivities in order to obtain better estimates of transmitted power (Piggott et al., 1957). Our current estimate of this power is based upon knowledge of fixed transmitter and receiver parameters. However, we believe that the seasonal and day to day variability of ground effects on radar power does not significantly affect our results. In fact, variability in transmitted power due to transmitter and antenna faults is a more serious problem. Even so, these can be accounted for within our forward model.

Despite these limitations, our retrieval makes use of what we believe to be the most complete model yet applied to ground based radar measurements of mesospheric electron densities. The forward model includes both absorption and Faraday rotation of each mode, parameterises partial reflections from small scale fine structure. It also supports reproduction of known profiles and supports the retrieval of $N_e(z)$ profiles which are consistent with measurements made by previous researcher, and, which support conventional ideas about wintertime transport of thermospheric NO to the mesosphere.

Acknowledgements

This work was undertaken as part of a Ph.D. project financed by the provision of a University of Canterbury Doctoral Scholarship.

References

Austin, G.L., 1971. A direct measuring differential phase experiment. Journal of Atmospheric and Terrestrial Physics 33, 1667–1674.

Belrose, F.F., Burke, M.J., 1964. Study of the lower ionosphere using partial reflection: 1. Experimental technique and method of analysis. Journal of Geophysical Research 69 (13), 2799–2818.

Belrose, J.S., 1970. Radio wave probing of the ionosphere by the partial reflection of radio waves (from heights below 100 km). Journal of Atmospheric and Terrestrial Physics 32, 567–596.

von Biel, H.A., 1977. The phase switched correlation polarimeter —a new approach to the partial reflection experiment. Journal of Atmospheric and Terrestrial Physics 39, 769–778.

von Biel, H.A., 1995. Ground based radar investigation of the Antarctic mesosphere. Advances in Space Research 16 (5), 91–98.

von Biel, H.A., 1999. Back-scattering of MF electromagnetic energy from the Antarctic stratosphere. Advances in Space Research 24, 1669–1677.

von Biel, H.A., Flood, W.A., Camnitz, H.G., 1970. Differential-phase partial-reflection technique for the determination of D-region ionization. Journal of Geophysical Research 75, 4863–4870.

Bilitza, D., 2001. International reference ionosphere 2000. Radio Science 36 (2), 261–275.

Budden, K.G., 1961. Radio Waves in the Ionosphere. Cambridge University Press, Cambridge.

Danilov, A.D., Rodevich, A.Y., Smirnova, N.V., 1995. Problems with incorporating a new D-region model into the IRI. Advances in Space Research 15 (2), 165–168.

Ferguson, B.G., 1981. Measurement of ionospheric winter anomaly in southern hemisphere. Nature 290, 485–487.

Fleming, E.L., Chandra, S., Barnett, J.J., Corney, M., 1990. COSPAR international reference atmosphere: 1986, Part II: middle atmosphere models, chapter 2: zonal mean temperature, pressure, zonal wind and geopotential height as functions of latitude. Advances in Space Research 10(12) 11–59.

Flood, W.A., 1968. Revised theory for partial reflection D-region measurements. Journal of Geophysical Research 73 (17), 5585–5598.

Fiedrich, M., Torkar, K.M., 2001. FIRI: a semiempirical model of the lower ionosphere. Journal of Geophysical Research 106 (A10), 21409–21418.

Garcia, R.R., Solomon, S., Avery, S.K., Reid, G.C., 1987. Transport of nitric oxide and the D-region winter anomaly. Journal of Geophysical Research 92 (D1), 977–994.

Gardner, F.F., Pawsey, J.L., 1953. Study of the ionospheric D-region using partial reflections. Journal of Atmospheric and Terrestrial Physics 3, 321–344.

Gregory, J.B., Manson, A.H., 1969. Seasonal variations of electron densities below 100 km at mid-latitudes, 1, electron densities and atmospheric circulation. Journal of Atmospheric and Terrestrial Physics 31, 682–701.

Hocking, W.K., 1987. Radar studies of small scale structure in the upper middle atmosphere and lower ionosphere. Advances in Space Research 7 (10), 327–338.

Hocking, W.K., Röttger, J., 1983. Pulse length dependence of radar signal strengths for Fresnel backscatter. Radio Science 18 (6), 1312–1324.

Hocking, W.K., Röttger, J., 1997. Studies of polar mesosphere summer echoes over EISCAT using calibrated signal strengths and statistical parameters, Radio Science 32 (4) 1425–1444.

Hocking, W.K., Vincent, R.A., 1982. Comparative observations of D region HF partial reflections at 2 and 6 MHz. Journal of Geophysical Research 87 (A9), 7615–7624.

Hocking, W.K., Fukao, S., Yamamoto, M., Tsuda, T., Kato, S., 1991. Viscosity waves and thermal-conduction waves as a cause of "specular" reflectors in radar studies of the atmosphere. Radio Science 26, 1281–1303.

Holt, O., Landmark, B., Lied, F., 1961. Observations of electron density and collision frequency during polar radio blackout conditions. Journal of Atmospheric and Terrestrial Physics 23, 318–329.

Kawahira, K., 1984. Dynamical influences of planetary waves on nitric oxide variations in the D-region. Journal of Atmospheric and Terrestrial Physics 46, 321–333.

Lawrence, B.N., Randel, W.J., 1996. Variability in the mesosphere observed by the Nimbus 6 pressure modulator radiometer. Journal of Geophysical Research 101 (D18), 23475–23489.

Mechtly, E.A., Bowhill, S.A., Smith, L.G., 1972. Changes of lower ionosphere electron concentrations with solar activity. Journal of Atmospheric and Terrestrial Physics 34, 1899–1907.

Piggott, W.R., Beynon, W.J.G., Brown, G.M., Little, C.G., 1957. The measurement of ionospheric absorption. Ann. IGY, III, Part II, Pergamon Press, London, Vol. 173.

Press, W.H., Teutolsky, S.A., Vettering, W.T., Flannery, B.P., 1992. Numerical Recipes in C, 2nd Edition. Cambridge University Press, Cambridge.

Rodgers, C.D., 1990. Characterization and error analysis of profiles retrieved from remote sounding measurements. Journal of Geophysical Research 95 (D5), 5587–5595.

Rodgers, C.D., 2000. Inverse Methods for Atmospheric Sounding: Theory and Practice. World Scientific, Singapore.

Schwentek, H., 1974. Wave-like structures in the variation of ionospheric absorption. Journal of Atmospheric and Terrestrial Physics 63, 1173–1178.

Sen, H.K., Wyller, A.A., 1960. On the generalization of the Appleton–Hartree magnetoionic formulas. Journal of Geophysical Research 65 (12), 3931–3949.

Siskind, D.E., Bacmeister, J.T., Summers, M.E., 1997. Two-dimensional model calculations of nitric oxide transport in the middle atmosphere and comparison with halogen occultation experiment data. Journal of Geophysical Research 102 (D3), 3527–3545.

Thomas, L., 1961. The winter anomaly in ionospheric absorption. Journal of Atmospheric and Terrestrial Physics 23, 301–317.

Thrane, E.V., Grandal, B., 1980. Observations of fine scale structure in the mesosphere and lower thermosphere. Journal of Atmospheric and Terrestrial Physics 43, 178–189.