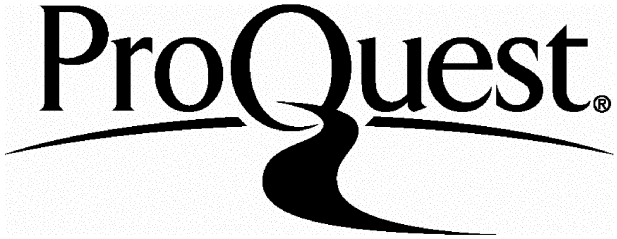
Comparison of high latitude thermospheric meridional neutral wind climatologies

Eoghan Michael Griffin

Department of Physics and Astronomy
University College London

Thesis presented for the degree of Doctor of Philosophy
of the University of London

July 2000



ABSTRACT

The combination of the long term databases of measurements from the Kiruna Fabry-Perot Interferometer and the EISCAT incoherent scatter radar, both covering more than a solar cycle of data, allows a unique comparison of the thermospheric meridional component of the neutral wind as observed by different experimental techniques. This allows the climatological behaviour of the neutral wind at high latitudes to be investigated, including the influence of both solar activity and season. Two techniques are applied to derive winds from the EISCAT database, one from previous work using the standard technique for incoherent scatter radars, and a new dataset derived using the Meridional Wind Model implementation of servo theory with the EISCAT data as input. The latter technique also uses contemporaneous EISCAT electric field data for correction to the equivalent servo winds.

Alongside the local measurements from experiment, model predictions of the behaviour of the winds can also be compared. These have been included and use both empirical sources as in the Horizontal Wind Model and Meridional Wind Model with International Reference Ionosphere input, and also the results from a first principles theoretical model, the UCL Coupled Thermosphere and Ionosphere Model.

Comparisons are made between the results from these techniques for each of eight categories corresponding to the four seasons, centred around the equinoxes and solstices, and for two solar activity levels. The detailed comparisons in each case and the implications of the results for the ability of the models to predict the long term behaviour of the winds and also for the degree of agreement between the techniques based on local measurements are discussed. Conclusions are drawn as to the major influences on the climatological behaviour of the wind at this latitude and the possibilities for further work to improve both experimental and modelling efforts.

Contents

Chapter 6 – Discussion

Appendix A: Theory of the Fabry Perot Interferometer

Chapter 1 - Introduction

1.1 Introduction

Research in Atmospheric Physics has developed in recent years by using many different approaches. Experimental techniques have been continually improved through the 20th century from the earliest observations of natural atmospheric emissions to current satellite-borne instrumentation giving global coverage and long-term datasets. The earliest observations were driven by curiosity about the naturally occurring aurora, and to this day one of the most challenging problems in atmospheric physics is to try to understand the complex situation at high latitudes.

Theoretical work has paralleled the experimental efforts, developing from early postulates about the nature and cause of the aurora to present day general circulation models (GCM's), which accurately describe the major features of the upper atmosphere composition and dynamics on a global basis.

In this thesis experiment and theory are used in tandem to accurately study the interaction of the neutral and ionized parts of the upper atmosphere.

1.2 Solar-Terrestrial System

The Earth is strongly influenced by the enormous energy output of the sun which is of the order of 3×10^{26} W (Akasofu , 1987). This output consists mainly of black-body radiation consistent with a solar temperature of 6000K, which penetrates the atmosphere as sunlight. The optical wavelengths provide the energy to heat the Earth itself and are not a primary source of energy for the upper atmosphere which they pass through.

Of more importance to upper atmospheric physics are the energy output from the sun in the form of X-ray, ultraviolet (UV) and extreme ultraviolet (EUV) electromagnetic radiation and the charged particles which are emitted by the sun. The electromagnetic radiation is absorbed, providing energy which contributes to the complex interactions between the constituent gases.

The flow of charged particles from the sun have a wide range of energies from less than 1 keV to more than 1 GeV. The lower energy portion of this spectrum below around 10 keV, which originates in the corona of the sun, is called the solar wind. This forms a plasma of equal numbers of positive and negative particles which are continually being ejected from the sun. Charged particles approaching the earth are

deviated by the geomagnetic field however they are able to reach the lower atmosphere within limited areas around the poles. This allows the solar wind to influence current systems, magnetic storms and auroral displays (C.T. Russell, 1987).

1.3 Atmospheric structure

The regions of the atmosphere have been named by words ending in "-sphere" describing either the chemical composition, temperature or dominant physical process within a certain range of altitude after the regime suggested by Chapman (Chapman, 1950). The upper boundary of a region represented in this manner uses the same word ending in "-pause" so that for the lowest region of the atmosphere, the *troposphere*, the upper boundary is called the *tropopause*. Descriptions of the atmosphere based on the vertical temperature structure, composition and degree of ionization follow.

1.3.1 Atmospheric Vertical Temperature Structure

The troposphere, stratosphere, mesosphere, thermosphere and exosphere are classifications of the atmosphere defined by the vertical temperature gradient, shown in Figure 1.1. The *troposphere* is influenced primarily by the heat radiated by the surface of the Earth which results in temperature decreasing with altitude to about 12 km with the temperature dropping by about 10 K per km. The *tropopause* is not fixed in altitude, nor are any of the other boundary regions, and the entire structure can vary to a large degree while retaining its essential height profile. As such Figure 1.1 is only representative of a profile for a "typical" terrestrial atmosphere.

For the altitude region between 20 and 50 km, known as the *stratosphere*, temperature rises with altitude. This is caused by the absorption of solar UV radiation in the range 200 to 300 nm by ozone (O_3) which is relatively abundant at these altitudes. Above the stratosphere, with the contribution from ozone absorption of UV radiation falling off, temperature drops with rising altitude to a minimum at around 85 km. This region is known as the *mesosphere* and its upper boundary is known as the *mesopause*. The mesopause is thought to be the coldest region of the atmosphere, with temperatures as low as 120 K reported (Roble et al., 1992).

Molecular oxygen will also absorb solar radiation with wavelengths lower than 175 nm during a process of dissociation into atomic oxygen. While the energy absorbed at low altitudes is much smaller than that absorbed by ozone, this source of heat becomes important at higher altitudes where the density is much reduced.

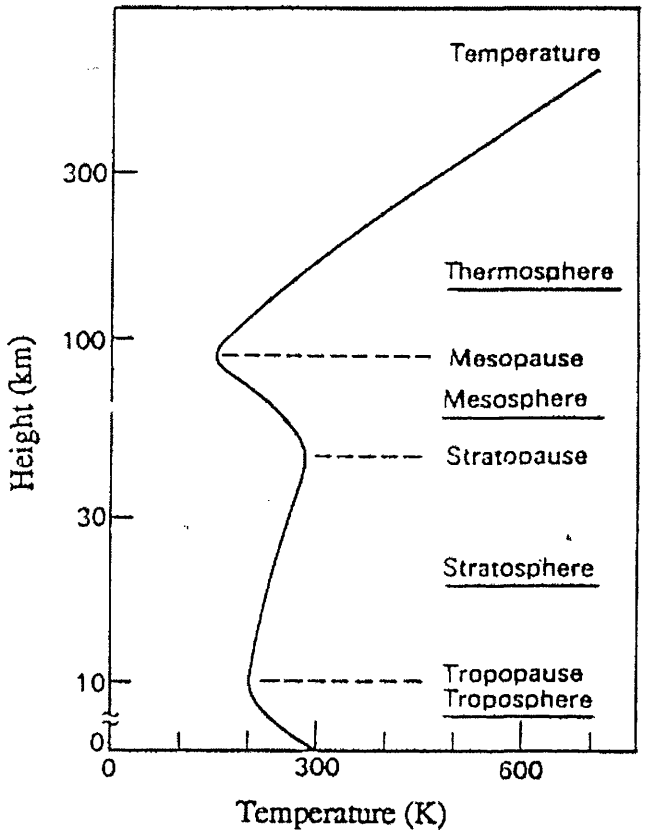


Figure 1.1 Atmospheric Temperature Structure, Nomenclature (after McEwan and Philips , 1975).

The region above the mesosphere is called the *thermosphere* and is characterised by a rise in temperature with altitude, which levels off to a maximum temperature, termed the exospheric temperature. The main heating process in the thermosphere is the absorption of solar EUV and UV radiation by the neutral gas and this causes a much higher temperature to be reached at these altitudes than at lower altitudes partly due to the much lower air density and the low rate of re-radiation. High thermal conductivity caused by the reduced pressure at altitudes typically around 300 km prevents any further increase in temperature. The actual temperature can range widely dependent on many variables, including for example solar cycle and seasonal conditions, from between 600 K to 1600 K.

Finally the formal structure of the atmosphere as described according to temperature gradient ends with the *exosphere*, a region which is characterised by the asymptotic maximum temperature attained at the top of the thermosphere.

The heat transport in the lower atmosphere is due primarily to radiative processes, however between 30 and 90 km the atmosphere is not in radiative equilibrium (Brasseur and Solomon, 1984). This is because in this region wave transience and dissipation are important and are a necessary inclusion for any model

calculations to correctly predict temperatures and winds. Between this region and the region dominated by thermal conductivity at high altitudes is a region in which convection is an important process. Heat may be transported from the lower thermosphere into the mesosphere by the process of turbulent mixing, also known as eddy diffusion. This provides a major heat loss mechanism for the thermosphere but only a minor heating process, compared to others, for the mesosphere (Hargreaves, 1979), a process which also contributes to the composition of the atmosphere.

1.3.2 Atmospheric Composition

The height variation of pressure and density within the atmosphere may be described by the hydrostatic equation, following the treatment of Hargreaves (1979).

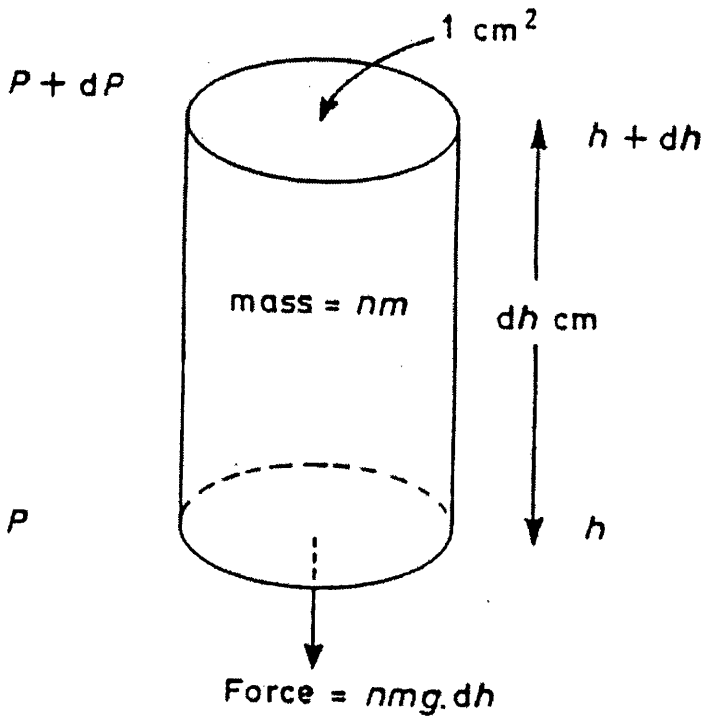


Figure 1.2 Forces on a cylinder of gas (after Hargreaves, 1979).

Figure 1.2 represents a gas containing n molecules per unit volume, each of mass m. If the height of a unit cross-section cylinder is dh then it contains total mass $nm\ dh$. In static equilibrium the net upward force due to the difference in pressure between the top and bottom faces balances the downward force experienced by the cylinder under gravity which is $nmg\ dh$, where g is the acceleration due to gravity:

$$(P + dP) - P = -nmg \cdot dh \qquad (1.1)$$

which gives

$$\frac{dP}{dh} = -nmg \tag{1.2}$$

By the gas law $P = nkT$, where k is Boltzmann's constant and T is the absolute temperature for the species. Hence

$$\frac{1}{P}\frac{dP}{dh} = -\frac{mg}{kT} = -\frac{1}{H} \tag{1.3}$$

where

$$H = \frac{kT}{mg} \tag{1.4}$$

is defined as the scale height. If Equation (1.3) is integrated, with H constant, then

$$P = P_0 \exp\left(-\frac{h}{H}\right) \tag{1.5}$$

where P_0 is the pressure at height $h = 0$. The value of H then gives the vertical distance in which P changes by a factor e. In the Earth's atmosphere H changes from about 5 km at 80 km to 70-80 km at 500 km. The scale height is an important parameter in specifying atmospheric properties because above about 100 km the constituents are not well-mixed. The principle constituent gases of the atmosphere are maintained in roughly constant abundances up to around 100 km by the process of turbulent mixing, even though the overall density decreases exponentially with altitude. This region is known as the *turbosphere* and its upper boundary is called the *turbopause*. Above the turbosphere the constituents are not well-mixed and composition changes with height according to the different scale heights of the constituent gases and this region is termed the *heterosphere*.

Within the heterosphere the atmospheric constituents separate out according to their relative masses under the influence of the gravitational field, the concentration of the lightest gases drop off more slowly with altitude due to molecular diffusion. As a result the lighter gases become relatively more abundant at higher altitudes as shown in Figure 1.3 which assumes an exospheric temperature of 800 K. The temperature of

the exosphere governs the altitudes at which composition changes between the major constituents.

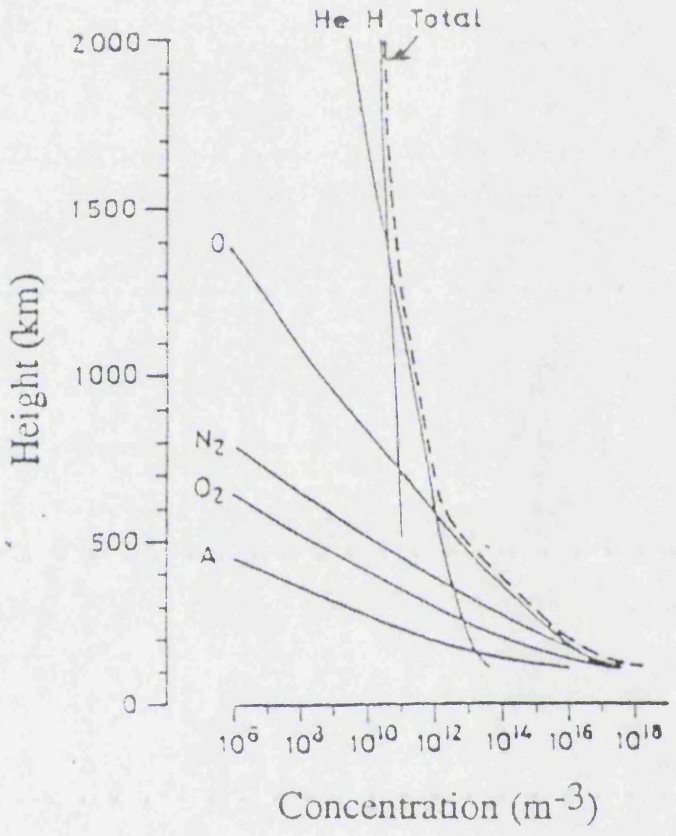


Figure 1.3 High Altitude Atmospheric Composition (after CIRA , 1972).

The composition of the atmosphere for the principle constituents at sea level is given in Table 1, and this is representative of the turbosphere as a whole. There are however many minor constituents within the turbosphere which play an important part in the physics of the atmosphere. Molecular oxygen is dissociated into atomic oxygen by solar UV radiation as follows

$$O_2 + h\upsilon \rightarrow O + O \tag{1.6}$$

where $h\upsilon$ is a quantum of radiation with wavelength below 175 nm. This production process leads to an increasing abundance of atomic oxygen above around 90 km, which matches molecular abundances at around 125 km and thereafter dominates over the molecular form.

Table 1.1 Atmospheric Composition at ground level (after Hargreaves , 1979).

Molecule	Mass (in a.m.u.)	Percent of Volume	Concentration / cm^3
Nitrogen	28.02	78.1	2.1×10^{19}
Oxygen	32.00	20.9	5.6×10^{18}
Argon	39.96	0.9	2.5×10^{17}
Carbon dioxide	44.02	0.03	8.9×10^{15}
Neon	20.17	0.002	4.9×10^{14}
Helium	4.00	0.0005	1.4×10^{14}
Water	18.02	Variable	

Another important minor constituent is ozone which is produced between 15 and 35 km as follows

$$O + O_2 + M \rightarrow O_3 + M \tag{1.7}$$

where M represents some third body which carries away excess energy. This reaction depends critically on the probability that the three atoms or molecules involved will collide at the same time and thus requires a relatively high pressure.

Also at these altitudes atomic oxygen is produced by radiation in the range 200 to 240 nm which is more penetrating than that causing dissociation at higher altitudes. Once produced the ozone itself may be dissociated by a wide range of radiation from 210 to 310 nm or may collide with an oxygen atom to produce two molecules of oxygen. The small concentration of ozone that remains has an important role to play in the heating of the atmosphere at stratospheric altitudes.

Another important minor constituent is nitric oxide (NO) which is produced in the mesosphere by the reaction

$$N + O_2 \rightarrow NO + O \tag{1.8}$$

where the atomic nitrogen in the reaction has been previously formed by other reactions. NO is then lost through a further reaction involving atomic nitrogen

$$NO + N \rightarrow N_2 + O \tag{1.9}$$

This results in minute concentrations of *NO* above 90 km but as will be seen later this minor constituent places an important role in the production of the lower ionosphere.

1.3.3 The Ionosphere

Solar X-ray and UV radiation together with energetic particles in the solar wind cause ionization within the atmosphere. This gives rise to the ionosphere, which is a classification of the atmosphere based on the vertical electron density profile, shown in Figure 1.4. There is considerable structure to this profile of electron density. The first region to be identified was called the E-region (E for Electric) at around 100 km. When the main peak was discovered it was called the F-region for alphabetical consistency. This region is sometimes subdivided into the F1-region (lower) and F2-region (upper) because a secondary peak below the main peak is often observed.

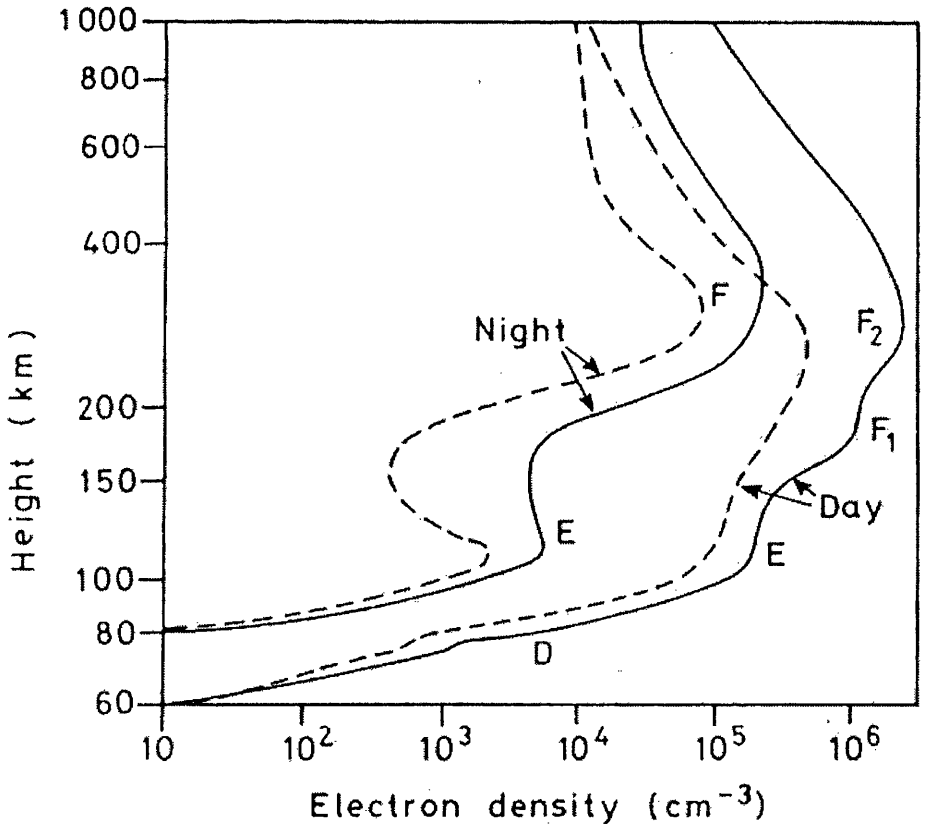


Figure 1.4 Vertical profile of electron density for the ionosphere (after Hargreaves, 1979).

Below the E-region exists a region of lower, highly variable, electron density. This is called the D-region extending from around 60 km to 90 km in altitude. The electron density throughout the ionosphere displays variations dependent on magnetospheric activity, strength of solar ionizing radiation, season, altitude and time of day.

Ionization is caused by different processes for different regions. Sources of ionization for the D-region include: X-rays of between 0.2 and 0.8 nm which will ionize all constituents, acting particularly on O_2 and N_2; ionization of NO by solar Lyman-α radiation at 121.5 nm; the portion of the solar EUV spectrum from 107.7 to 111.8 nm which will ionize excited oxygen molecules; cosmic rays also cause ionization and these will also ionize all constituents.

In the E-region ionization is due to solar EUV radiation in the 80 nm to 102.6 nm range which is absorbed by molecular oxygen to form O_2^+, and X-rays from 1 to 10 nm. X-rays at these wavelengths are able to ionize all constituents present in the E-region, forming in particular N_2^+, O_2^+ and O^+. The contribution of ionization due to X-rays is however heavily dependent on the solar cycle, being greatly reduced at solar minimum. As a result of subsequent reactions the most abundant ions in the E-region are NO^+ and O_2^+.

In the F1-region solar EUV radiation in the range 20 nm to 91 nm is believed to give rise to ionization with the 50 nm to 60 nm portion of the radiation being most heavily absorbed. Production of O_2^+, N_2^+, O^+, He^+ and N^+ due to primary ionization eventually result through reactions in NO^+ and O_2^+ being the most abundant ions in the F1-region, similar to the case in the E-region.

By day NO^+ and O_2^+ predominate below 165 km with O^+ being dominant above this height. At night this transition from molecular ion dominance to atomic ion dominance is found at about 220 km (Rishbeth and Garriott, 1969).

The distinction between the F1 and F2-regions is not drawn due to changes in the processes giving rise to electron production. To understand how the peak in the F2-region comes about it is necessary to examine the processes of charge exchange and rearrangement occurring in the ionosphere.

Electron recombination processes can be of three types :

Radiative recombination : $e + X^+ \rightarrow X + h\upsilon$

Dissociative recombination : $e + XY^+ \rightarrow X + Y$

Attachment : $e + Z \rightarrow Z^-$

Due to the low pressures in the E and F-regions attachment is not an important mechanism in these regions. The recombination coefficients are 10^{-13} and 10^{-18} m^3s^{-1} for the dissociative and radiative processes respectively (Hargreaves, 1979). The dissociative process is much quicker than the radiative process as energy and momentum are more easily conserved if two bodies are produced. In the situation where molecular ions are present, which is the case in the E-region, the dissociative recombination occurs in the following manner:

$$e + O_2^+ \rightarrow O + O$$

$$e + N_2^+ \rightarrow N + N$$

$$e + NO^+ \rightarrow N + O$$

In the case where electrons recombine directly with positive ions the rate of recombination can be written as:

$$L = \alpha [X^+] N_e = \alpha N_e^2$$

where the square brackets denote concentration, N_e is the electron density and if ionisation is dominated by one species as it is in the F2 region, $[X^+] = N_e$ so that within the medium charge neutrality is maintained. α is the recombination coefficient.

The attachment method can be represented using the following equation to describe the rate of attachment in terms of N_e and β which is the attachment coefficient.

$$L = \beta N_e$$

For the F-region electron loss occurs in a two-stage process with two possibilities for this process:

(a) $O^+ + N_2 \rightarrow NO^+ + N$

$\Downarrow$

(b) $NO^+ + e \rightarrow N + O$

or

(a) $O^+ + O_2 \rightarrow O_2^+ + O$

$\Downarrow$

(b) $O_2^+ + e \rightarrow O + O$

For (a) above the reaction rate is $\beta[N_2]$ and for (b) the reaction rate is $\alpha[NO^+]N_e$. For low altitudes β is large so that O^+ is rapidly converted to NO^+. The overall loss rate is controlled by (b) and α-type processes should dominate. At high altitudes β is small, leading to (a) dominating the overall reaction rate. In this manner behaviour changes from α-type to β-type as height increases.

To see how this affects the formation of the F1 and F2-regions consider that the electron production rate, q, is proportional to $[O]$ as most of the F-region ionization is derived from atomic oxygen. Also β is altitude dependent as a result of its dependence on the concentration of neutral molecules, $[N_2]$.

For the equilibrium electron density N_e then

$$N_e = \frac{q}{\beta} \propto \frac{[O]}{[N_2]} \propto \exp\left[-\frac{h}{H(O)} + \frac{h}{H(N_2)}\right] = \exp\left[0.75\frac{h}{H(O)}\right] \tag{1.9}$$

which follows from the masses of N_2 and O being in the ratio 1.75:1, with $H(O)$ and $H(N_2)$ the scale heights of the respective species at a height h.

The α-type processes would, if acting alone, give rise to a peak in electron density centred on the peak of electron production. The β-type processes alone would give a steady rise in electron density with height. The consequence of the addition of the two processes is that, for the F-region, if the height at which the transition to β-type from α-type processes is below the height of peak production, then no secondary F1 peak is formed.

If however the reverse is true and the transition height is above the height of peak electron production, then above the production height the electron density increases as β decreases and the F1 region is formed.

The electron density drops above the F2 peak however as at these altitudes the atmospheric density drops off and chemical recombination becomes less important as a loss mechanism than vertical transport. Vertical transport depends on the diffusion coefficient which changes with height so that vertical transport becomes more efficient with increasing altitude.

1.4 Atmospheric Emissions

The photochemical reactions which occur in the upper atmosphere can produce radiation in the ultraviolet, visible or infra-red regions of the spectrum. The resultant emissions are called the airglow. These emissions depend on the creation of an excited species and its subsequent return to its ground state. Apart from the solar radiation the causes of such excitations can be radiative recombination, excited reaction products and ionization processes.

The airglow includes both dayglow and nightglow. The nightglow is by far the more widely studied due to the difficulty associated with studying the dayglow, it being very difficult to isolate the emission of interest from the intense background of sunlight scattered by air molecules, aerosol and dust particles. The airglow is usually weak in intensity, at subvisual levels, but continuous in emission and not confined to specific latitudes.

The aurora also contributes to atmospheric emissions and is caused by protons and electrons entering the Earth's magnetic field and following the magnetic lines of force down into the atmosphere at high latitude. Here they collide with the constituent species and produce radiative emissions. The aurora is therefore usually confined to high latitudes and is discontinuous in emission, but often is intense enough to be visible to the naked eye.

Both airglow and aurora result from the excitation of the constituent atmospheric species such as N_2, O_2, O, N, H and OH, but the airglow is not primarily caused by the energetic particle mechanism, as for the aurora.

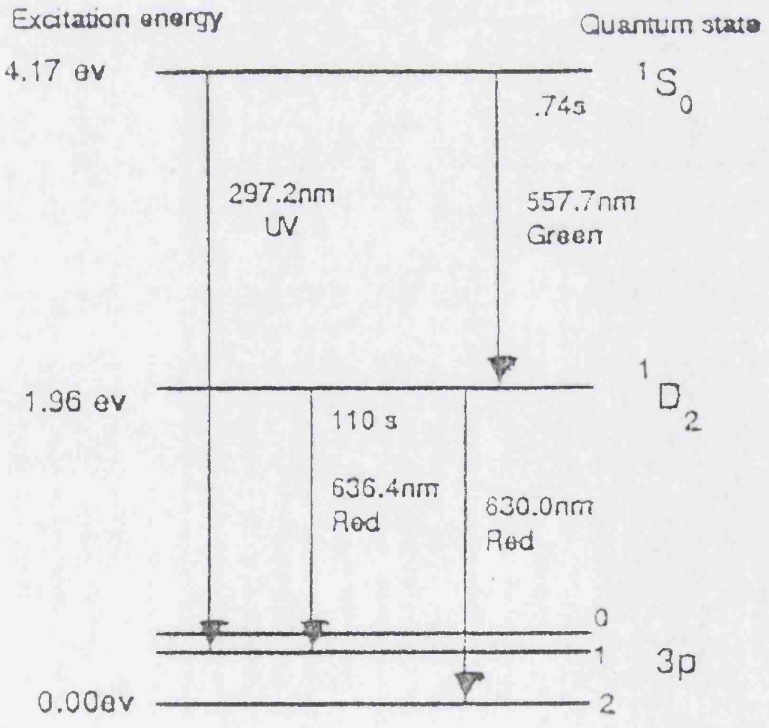


Figure 1.5 Energy Level Diagram for Atomic Oxygen (after Wayne , 1985).

1.4.1 Atomic Oxygen Emissions

One of the most prominent of the excited species contributing to the airglow is atomic oxygen (OI) which produces its dominant optical emission lines at 630 nm and 577.7 nm. Figure 1.5 shows the energy level diagram for the lower excited states of atomic oxygen showing the 630 nm line to originate from the decay of the atom from the excited O(^{1}D) state, with a 110 second lifetime, to the ground state O(^{3}P). The 577.7 nm line originates from the decay of the atom from the metastable O(^{1}S) state, with a lifetime of 0.74 seconds, to the O(^{1}D) state.

Table 1.2 Atomic Oxygen Production and Loss Mechanisms (after Witasse et al., 1999). In this table e_{ph} indicates a photo-electron, e_{th} an thermal electron

$O(^1D)$	
Production	Loss
$O + e_{ph} \rightarrow O(^1D) + e_{ph}$	$O(^1D) + N_2 \rightarrow O + N_2$
$O + e_{th} \rightarrow O(^1D) + e_{th}$	$O(^1D) + O_2 \rightarrow O + O_2$
$O_2^+ + e_{th} \rightarrow O + O(^1D)$	$O(^1D) + O \rightarrow O + O$
$O_2 + hv \rightarrow O + O(^1D)$	$O(^1D) + e_{th} \rightarrow O + e_{th}$
$O(^1S) \rightarrow O(^1D) + hv(557.7\ nm)$	$O(^1D) \rightarrow O + hv(630.0\ nm)$
$N(^2D) + O_2 \rightarrow NO + O(^1D)$	$O(^1D) \rightarrow O + hv(636.4\ nm)$
$N^+ + O_2 \rightarrow NO^+ + O(^1D)$	

$O(^1S)$	
Production	Loss
$O + e_{ph} \rightarrow O(^1S) + e_{ph}$	$O(^1S) + O_2 \rightarrow O + O_2$
$O_2^+ + e_{th} \rightarrow O + O(^1S)$	$O(^1S) + O \rightarrow O + O$
$O_2 + hv \rightarrow O + O(^1S)$	$O(^1S) \rightarrow O + hv(557.7\ nm)$
$N + O_2^+ \rightarrow NO^+ + O(^1S)$	$O(^1S) \rightarrow O + hv(297.2\ nm)$
$O_2 + O \rightarrow O_2 + O(^1S)$	
$N_2(A^3\Sigma_u^+) + O \rightarrow N_2 + O(^1S)$	

Table 1.2 shows the reaction processes identified by Witasse et al. (1999) which give rise to the O(^{1}D) and O(^{1}S) excitations and their associated loss mechanisms. These reactions represent production through electron impact, dissociative recombination, photodissociation of O_2 and chemistry for both

emissions. Additionally O(^{1}S) may be produced by collisional deactivation of N_2 and O(^{1}D) can be produced by cascade from O(^{1}S). For both emissions the loss mechanisms identified represent collisional deactivation and radiative transition.

Examining the collisional deactivation (quenching) loss process, Roach and Gordon (1973) assumed a steady state condition and wrote

$$P(h)=\frac{N_y}{\tau_y}+k_q N_y N_z \tag{1.10}$$

where $P(h)$ is the production rate of the excitations of interest, N_y is the number density of the emitting species, τ_y is the lifetime of the excited state, N_z is the number density of the quenching species and k_q is the quenching rate constant.

Further assuming that the emission of interest is the only one from the species N_y they obtained

$$E(h)=\frac{P(h)}{1+k_q N_z \tau_y} \tag{1.11}$$

where $E(h)$ is the emission rate. Thus when quenching is negligible, production and emission rates are equal. This however is never the case for atomic oxygen in the thermosphere, quenching being an important mechanism as shown by Figure 1.6, which is based on a quenching rate of 10^{-13} cm^{3}s^{-1} collisions and also features the emission rates for the O(^{1}S) and O(^{1}D) excitations at heights of 100 km and 140 km respectively.

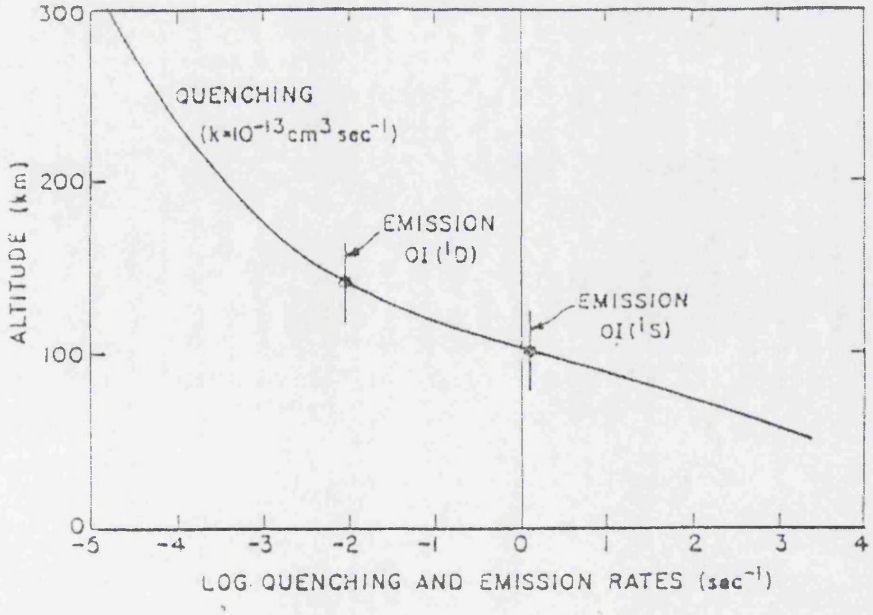


Figure 1.6 Collisional De-excitation rate in the Upper Atmosphere as a function of altitude (after Roach and Gordon, 1973).

The emission line at 630 nm is faint at 100 km as the quenching rate is such that O(^{1}D) atoms are de-excited by collision before radiation occurs. Indeed quenching of the O(^{1}D) state by N_2 and O_2 below around 200 km leads to no appreciable emission, while at higher altitudes of above about 350 km the rapid decrease in O_2 concentrations diminishes formation of the recombining molecular ions, and therefore the O(^{1}D) emission (Sipler et al., 1982).

The emission rate shows therefore the lower limit at which emission can occur for the 630 nm line and similarly in the case of the O(^{1}S) atoms giving rise to the 577.7 nm line. This agrees with observations of 577.7 nm at 90 to 110 km and 630 nm only at greater heights (Roach and Gordon, 1973), most of the emission originating from a typical altitude of 250 km (Hernandez and Roble, 1976). Figure 1.7 shows the altitude profiles resulting from the balance of the production and quenching mechanisms mentioned.

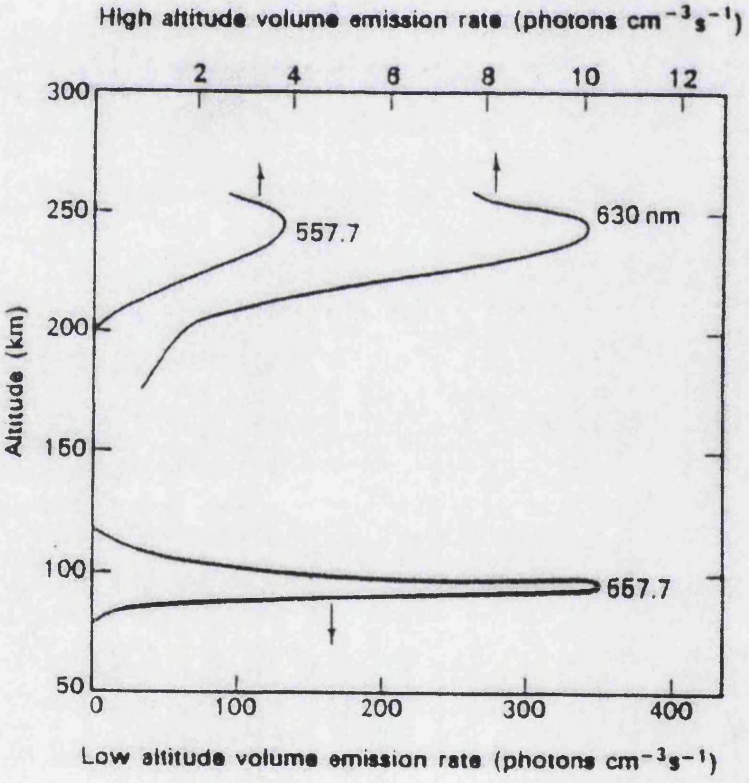


Figure 1.7 Altitude profiles for red and green atomic oxygen lines in the nightglow (after Gulledge et al., 1968)

With a radiative lifetime of 110 seconds the O(^{1}D) emission at 630 nm may be measured to determine the properties of the layer of the atmosphere at which they

originate as the mean time between collisions at 250 km is 0.3 seconds. As this time is short compared to the radiative lifetime of the emission we can assume that the emitting atoms have had time to thermalize before emission. Thus, we can use the emission as a tracer which is representative of the population as a whole in the region of emission.

The O(^{1}S) emission has a radiative lifetime of only 0.74 seconds but because of the increased air density at the lower altitude to which the peak of this emission is assigned the mean time between collisions is much shorter than that at the higher altitude for the O(^{1}D) emission. This allows the assumption that the emitting atoms have had time to thermalize before emission to be used for this emission also.

1.5 Thermospheric Circulation

The rotation of the earth about its axis causes the absorption of solar radiation within the upper atmosphere to vary on a daily basis. Heating on the dayside of the Earth causes expansion of the atmosphere while on the nightside the atmosphere contracts. The resultant heating pattern produces pressure gradients which drive a global thermospheric circulation. The pressure gradients minimize the day-night differences in pressure and temperature by transporting heat from the warm dayside to the cool nightside (Kohl and King, 1967).

To describe the thermospheric horizontal momentum for a neutral gas a reasonably complete equation which treats the atmosphere as a fluid can be written as follows (Rishbeth and Garriot, 1969),

$$2\vec{\Omega}\times\vec{U}+\frac{\partial\vec{U}}{\partial t}+\left(\vec{U}\nabla\right)\vec{U}=\vec{g}-\frac{1}{\rho}\nabla P-\nabla\Psi+\left(\frac{\mu}{\rho}\right)\nabla^2\vec{U}-\upsilon_{in}\left(\vec{U}-\vec{V_i}\right) \qquad (1.13)$$

where $\vec{U}$ is the wind vector, $\vec{V_i}$ is the ion drift vector, $\vec{\Omega}$ is the angular momentum vector for the Earth, g is the acceleration due to gravity, ρ is the mass density, P is the pressure, μ is the coefficient of viscosity, υ_{in} is the ion-neutral collision frequency and Ψ is the tide raising potential.

Equation (1.13) can be read as

Coriolis Force + Acceleration + Momentum Transport =

Gravity - Pressure Gradient - Tidal Potential + Viscous Drag - Ion Drag

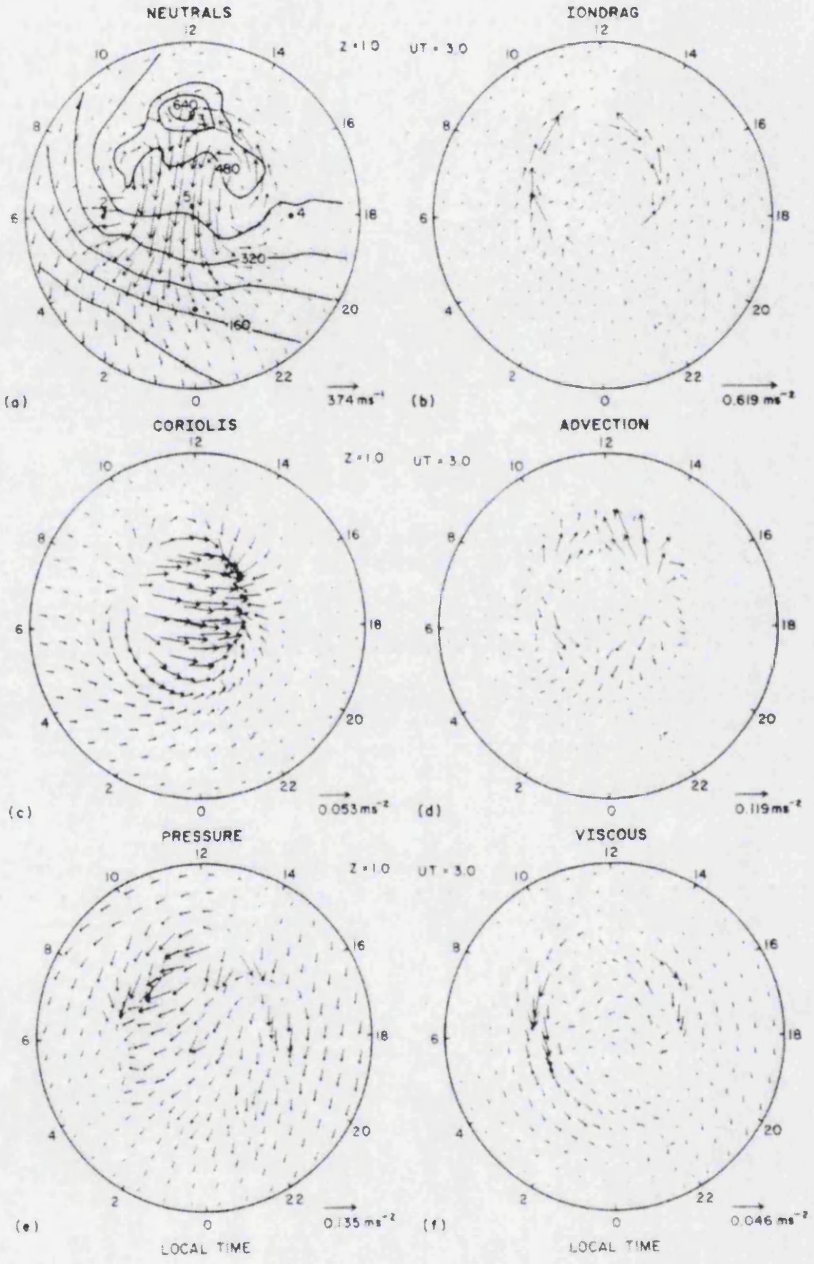


Figure 1.8 Modelled momentum forces in the thermosphere at 0300 UT for the southern polar cap. Plot (a) represents the resultant neutral winds, the advection force (Plot (d)) refers to the momentum transport term in Equation (1.13) (after Killeen and Roble, 1984).

The Coriolis force acts on a volume of neutral gas (parcel) in motion causing it to gain a component of velocity perpendicular to its direction of motion. The momentum transport term describes the acceleration caused due to position within a velocity field. The pressure-gradient force causes motion opposite to the direction of

the pressure gradient. The viscosity term is due to the atmosphere, considered as a fluid, having a resistance to shear as it begins to move in response to the driving forces. The ion drag term refers to the effect whereby the neutral winds often have a higher velocity than the ions, thus the ions tend to reduce the velocity of the neutral gas through collisions. It is also possible in highly disturbed geomagnetic conditions for the ions to accelerate the neutrals through collisions. The ion drag and pressure gradient terms combine to initiate motion for the parcel of neutral gas. Figure 1.8 shows output from a thermospheric numerical model demonstrating the major momentum terms seperately and also the resultant wind flow pattern, as they are represented at the southern polar cap.

Large amplitude, vertically propagating waves are a feature of the upper atmosphere and these waves can have periods of from several minutes to one day. The vertical propagation of these waves leads to an exponential increase in their amplitudes as they move into regions with lower air density. Waves called thermal tides are caused by the daily variation in the solar UV radiation absorption of ozone in the stratosphere. These thermal tides have periods of 12 and 24 hours and other harmonics of a diurnal forcing mechanism and can produce winds of 100 m s^{-1} as they propagate towards the upper atmosphere.

Seasonal variations also influence the global circulation as the maximum solar heating occurs at different latitudes. For the equinoxes the maximum is in the equatorial region. This produces winds flowing towards the cooler polar regions from the hotter and higher pressure equatorial regions. At solstices the maximum heating occurs in the summer hemisphere, minimising in the winter hemisphere. The resultant asymmetry provides a summer to winter circulation which causes a summer to winter hemisphere horizontal wind and upward motion over the summer hemisphere and downward motion over the winter hemisphere.

In the absence of electric fields, ionization can move freely only in a direction parallel to the Earth's magnetic field. Thus, the pressure exerted by a horizontal, equatorward wind in the neutral atmosphere forces ionization up the magnetic field lines, while a poleward wind carries the ionization to lower heights. Since they produce large changes in the effective recombination rate of the ionization, it is clear that changing meridional winds, often associated with a change in magnetic declination, are the major cause of most changes in the F region (Titheridge, 1993). This includes the large variations observed at different seasons, at different longitudes and between the two hemispheres.

At low latitudes, electric fields are produced by zonal winds driving E region ionization across the magnetic field lines. These fields map up into the F region to produce vertical drifts of ionization, which are responsible for most features of the

equatorial anomaly. Meridional winds play some part in producing an inter-hemispheric flow of ionization, which can alter the size of the equatorial anomaly peaks (Titheridge, 1995b).

Large electric fields are produced in the magnetosphere. These map down to the ionosphere at high latitudes where they cause a large horizontal convection of the field lines and the accompanying plasma. Thus, in the polar regions electric fields tend to drive the winds rather than dragging the winds as in the mid- and low latitudes (Killeen et al., 1995). Some of the magnetospheric electric field also penetrates to mid-latitudes, causing ionization drifts which are normally small and less important than movements driven by neutral wind.

Chapter 2 - Measurement Techniques and Models of the Upper Atmosphere

2.1 Introduction

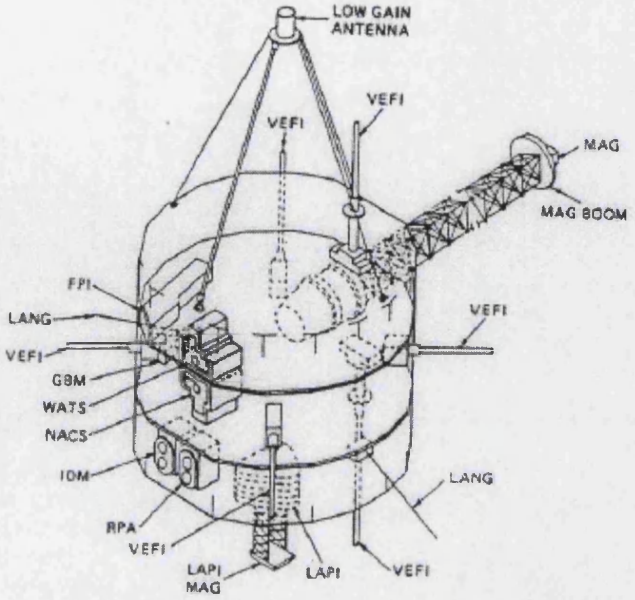
The development of geoscience in general and atmospheric research in particular has been driven from the earliest experiments by the development of instrumentation and measurement techniques. In more recent times the experimental efforts have been complemented by increasingly sophisticated models of the atmosphere, the development of which has been assisted by the parallel development and availability of computing resources. This chapter details the measurement techniques used in the study of the upper atmosphere. The development of atmospheric models is also introduced and the relevance of these models to the subject of ion-neutral coupling is discussed. Finally the most important specific experiments relating to ion-neutral coupling are discussed.

2.2 Neutral Atmosphere Measurement Techniques

Many different techniques have been developed in order to study the atmosphere and determine its physical and chemical properties. The methods employed reflect the most important parameters needed for a better understanding of the atmosphere. These include studying the constituents of the atmosphere and their relative abundances, the temperatures of different regions and the winds operating on the constituents.

As methods of atmospheric observation have advanced, the emphasis has moved from identification of emission features and their excitation processes to quantification. With the combination of ground-based and in-situ measurements using optical and radio methods it is now possible that the "weather" and "climate" of the thermosphere/ionosphere system can be monitored. This depends to a large extent on the use of accurate theoretical models of the parameters being measured to establish "normal" conditions and also detailed and rigourous understanding of the physical and chemical processes involved in the emission and transfer of radiation (Solomon, 1991).

The techniques developed in order to establish a database for these atmospheric parameters fall broadly into three categories: direct sensing, indirect sensing and remote sensing.



DYNAMICS EXPLORER INSTRUMENT	MAIN MEASUREMENT
Fabry-Perot interferometer (FPI)	meridional neutral wind (u_n) and neutral temperature (T_n)
Wind and temperature spectrometer (WATS)	zonal (v_n) and vertical neutral wind (w_n) and temperaure (T_n)
Retarding potential analyzer (RPA)	meridional ion drift (u_i) and ion temperature (T_i)
Ion drift meter (IDM)	zonal (v_i) and vertical ion drift (w_i)
Neural mass spectrometer (NACS)	neutral constituent number densities of N_2, O, N, Ar, and He
Langmuir probe (LANG)	electron (N_e) and ion (N_i) number density and electron temperature (T_e)
Low altitude plasma instrument (LAPI)	precipitating plasma fluxes and pitch angle distribution
Vector electric field instrument (VEFI)	electric field (E)
Magnetic field instrument (MAG-B)	magnetic field (B)
Galactic background monitor (GBM)	star sensor for precise orientation of measurements

Figure 2.1 The Dynamics Explorer 2 satellite (after Killeen and Roble, 1988)

2.2.1 Direct Sensing

Direct sensing involves using an instrument to take measurements from within the region of the atmosphere under investigation. Instruments are usually carried by rocket or satellite. Satellite platforms have provided great scope for atmospheric measurements over long periods, with full global coverage. Polar orbiting satellites have been launched such as Nimbus and Tiros which made up to 14 orbits in a 24 hour period and provided full global coverage at least twice per day (Houghton, 1977).

The Dynamics Explorer 2 satellite, shown diagramatically in Figure 2.1, was launched in 1983 and carried several instruments for direct sensing of the upper atmosphere. This mission produced a large database of measurements of atmospheric variables, enabling improvements to be made to the existing models of the atmosphere. Instruments on board included a mass spectrometer, a Fabry-Perot interferometer and a Langmuir probe.

The mass spectrometer is an instrument which can be used to determine the masses of ions or neutrals in the atmosphere. The decrease in air density at higher altitudes allows experiments to be carried out in conditions approximating a vacuum, and in the case of the mass spectrometer this is important as the principle of its operation is to measure the effect of electric and magnetic fields on the free motion of ions. In this respect it is necessary to ionize neutral constituents by electron bombardment prior to their entry into the apparatus.

The Langmuir probe is a device which is placed in a plasma and from the characteristic curve of current versus voltage for the probe, conclusions can be drawn on the conditions within the plasma. In regard to atmospheric research it is possible to determine the electron temperature for regions of the ionosphere and the relative abundances of the most important ions in these regions.

The Fabry-Perot interferometer is a high resolution spectroscopic device which is based on the principle of optical interference. It is used in atmospheric studies primarily due to its large light gathering power combined with high resolution. The range of instruments on board the satellite demonstrates the usefulness of satellite-based direct sensing, as a dataset can be built up consisting of measurements taken at the same altitude and position, from many different instruments.

2.2.2 Indirect Sensing

Indirect sensing has long been used as a method for atmospheric research, where measurements are taken, from outside the region of interest, of the characteristics of a sensor placed within the region. Satellites in near-earth orbits (200-1100 km) are influenced by drag, which produces a decay in their orbits due to the frictional forces experienced as they traverse the upper atmosphere. This drag may be

deduced by tracking the orbits of satellites. The density of the atmosphere at altitudes of hundreds of kilometers may be inferred using this technique, which has been in use since satellites were first launched in the late 1950's.

Rockets are also used for indirect sensing, sometimes by dropping spheres whose rate of descent is measured, and similarly the density of the atmosphere can be inferred by the effect of drag. This method is most useful between 90 and 130 km, below the altitude at which satellite orbits can be maintained. Luminescent chemicals are used at altitudes above 80 km also, providing wind data through measurements of the trail produced by the chemicals. This technique is usually restricted to nighttime and twilight though sodium trails have been used in daytime. At lower altitudes balloons have been used to take similar measurements since the 1940's.

2.2.3 Remote Sensing

The technique of remote sensing is based on either, (i), using a transmitter to propagate electromagnetic radiation through the region of the atmosphere being investigated and detecting the scattered radiation, or (ii), studying the radiation emitted from within the region of interest by the constituent species. The first method is called active remote sensing: examples include the incoherent scatter technique, which will be discussed in Section 2.3.1, and light detection and ranging (LIDAR). The second method is called passive remote sensing as it uses naturally occurring processes in the atmosphere as the source of its measurements. The most common type of passive remote sensing technique is to observe naturally occurring emissions from the constituent species within the atmosphere, at visible or infrared wavelengths.

LIDAR works on the principle of scattering light at optical wavelengths and recording a plot of intensity versus time lag, between emission of a signal from a laser and its detection. This instrument can be viewed as an optical analogue of a radar. A variety of scattering processes combine to produced the backscattered spectrum and a number of techniques have been developed to take advantage of these. Inelastic scattering processes may be resonant, Raman, resonant-Raman and fluorescence. Elastic scattering processes are Rayleigh and Mie.

In a Rayleigh Lidar the Rayleigh-backscattered spectrum from air molecules is recorded and the vertical density and temperature inferred from the profile. This process is useful in the middle atmosphere (from 30 km to around 90 km) but is limited at lower altitudes by the presence of aerosols which enhance the contribution of Mie scattering and at higher altitudes by the decrease in returned signal, which falls off as the atmospheric density rapidly decreases with height (Chanin and Hauchecorne, 1991).

Passive remote sensing commonly involves using interferometers, dispersion spectrometers or photometers to study the naturally occurring atmospheric emissions. The use of photometers allows measurement of photon intensity at selectable wavelengths using narrow bandpass filters (0.3 - 1 nm) to isolate a portion of the spectrum. Dispersion spectrometers use a different method whereby light from a broadband source such as the night sky is incident on a dispersion apparatus, such as a ruled grating, and is separated into its constituent wavelengths which may then be detected individually. Using this method allows the measurement of light intensities over a larger wavelength range.

Interferometers, such as the Fabry-Perot and Michelson interferometer, divide incident radiation into a number of coherent rays which are recombined to produce optical interference. This method provides very high spectral resolution and has been used to study atmospheric emissions from the earliest days of research in the field (Babcock, 1923). This thesis involves the use of Fabry-Perot interferometers to make observations of atomic oxygen emissions by passive remote sensing and details of the method and relevant implementation will be given in Chapter 3.

2.3 Ionospheric Measurement Techniques

By far the most widely available instrument for ionospheric research is the ionosonde which uses radio waves to determine the vertical electron density profile of the ionosphere. These instruments have been developed over many decades, having been first used in the 1930's, and will be described in detail later in this section. A number of other techniques have also been developed to allow more of the plasma parameters associated with the ionosphere to be measured. The most prominent instruments based on these techniques have been the incoherent and coherent scatter radars.

All forms of atmospheric scatter radar detect echoes scattered from within a medium where spatial irregularities in its refractive index are present, irrespective of the medium being ionized or not. By standard convention, radars that detect the thermal fluctuations of a medium, where the coherence time is short, have been termed incoherent-scatter radars, while those that scatter off physical structures within the medium, which tend to vary more slowly leading to long coherence times, are called coherent-scatter radars (Hargreaves, 1992). In reality, all scattering radars are to some extent 'coherent". Those labelled "incoherent" may properly be called quasi-coherent.

2.3.1 Incoherent Scatter

The incoherent scatter technique is based on the work of J.J. Thomson who showed that the power re-radiated by an electron which has been caused to oscillate by an electromagnetic wave is equal to the power in the incident wave that would fall on an area, called the scattering cross-section of the electron, given by

$$\sigma = \frac{\pi \mu_0^2 e^4}{m^2} \tag{2.1}$$

where μ_0 is the permeability of free space. Using a wave of power P incident on a volume V at a height h which has electron density n, the total returned signal has been shown to be (Ratcliffe, 1972)

$$\frac{P\sigma nV}{4\pi h^2} \tag{2.2}$$

By measuring the returned signal from a wave of known power, n may be determined. Radio waves are used for these measurements and because of the small scattering cross-section for the electron, about 10^{-28} m^2, very large antennae are necessary to pick up the returned signal.

The development of post-war radars led to the suggestion that they could be used to detect echoes from electrons in the upper atmosphere (Gordon 1958). The expected returned signal would be very weak because of the small scattering cross-section for the electron. The returned frequency band was also expected to be spread over a wide range of several MHz, equivalent to the Doppler spread of the electrons' thermal motions.

The first incoherent scatter experiments were carried out by Bowles in 1958. While the returned signal power was as expected, the shape and range of the returned frequency spread was very different from that expected. A typical returned spectrum for incoherent scatter radar is shown in Figure 2.2. The spread is only a few tens of kHz and there is a characteristic double hump shape.

This result was rapidly explained using different approaches by a number of authors (Fejer 1960, Dougherty and Farley 1960, Hagfors 1961), whose conclusions were similar. The characteristic shape reflects the state of the ions in the plasma from which the signal has returned rather than the electrons from which the reflection actually occurs. This is because the ions control the motion of the electrons. The reflected wave is really the result of constructive interference between the incident radiation and ion-acoustic waves in the plasma. Although they are ion-acoustic waves, oscillations of the same frequency will be impressed on the electrons by the ion motions. The central double-humped spectrum is caused by interference with the ion-

acoustic waves and the two outer spikes are due to interference with electron acoustic or plasma waves.

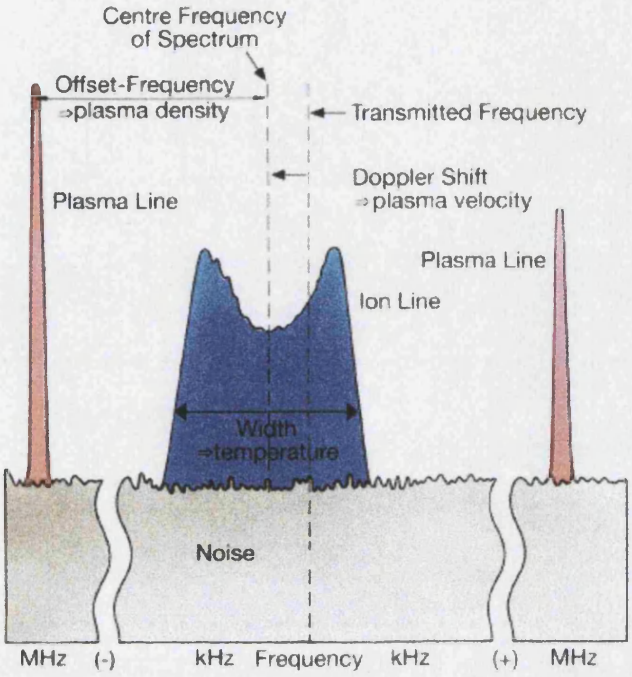


Figure 2.2 Returned spectrum from incoherent scatter

From the returned spectrum a number of important ionospheric parameters may be determined such as electron density, electron temperature, ion temperature, line-of-sight velocity and ion composition.

2.3.2 Coherent Scatter

Coherent scatter of a pulsed signal will produce an echo from which the signal from a particular range cell may be extracted by "gating" according to the time delay. The intensity of this returned signal will be related to the scattering mechanism and the mean frequency may show a Doppler shift according to the line-of-sight velocity of the scattering irregularities. The shape of the returned signal spectrum may also contain information about both the scattering mechanism and the scattering medium.

Coherent scatter radars can be used to study irregularities in both the E and F regions of the ionosphere. In the E region VHF radars may be used to detect irregularities associated with activity in the auroral electrojet which generates instabilities by the two-stream and gradient-drift mechanisms. The waves reflected from the electrojet travel almost perpendicular to the geomagnetic field and hence these field-aligned irregularities lead to the radar echoes being strongly aspect sensitive. To obtain the strongest echoes, the irregularities must be illuminated in a

direction perpendicular to the local field line. Similar instabilities form in the equatorial region, due to the equatorial electrojet. At these low latitudes, the geomagnetic field is almost horizontal to the Earth's surface and so the radar is pointed vertically.

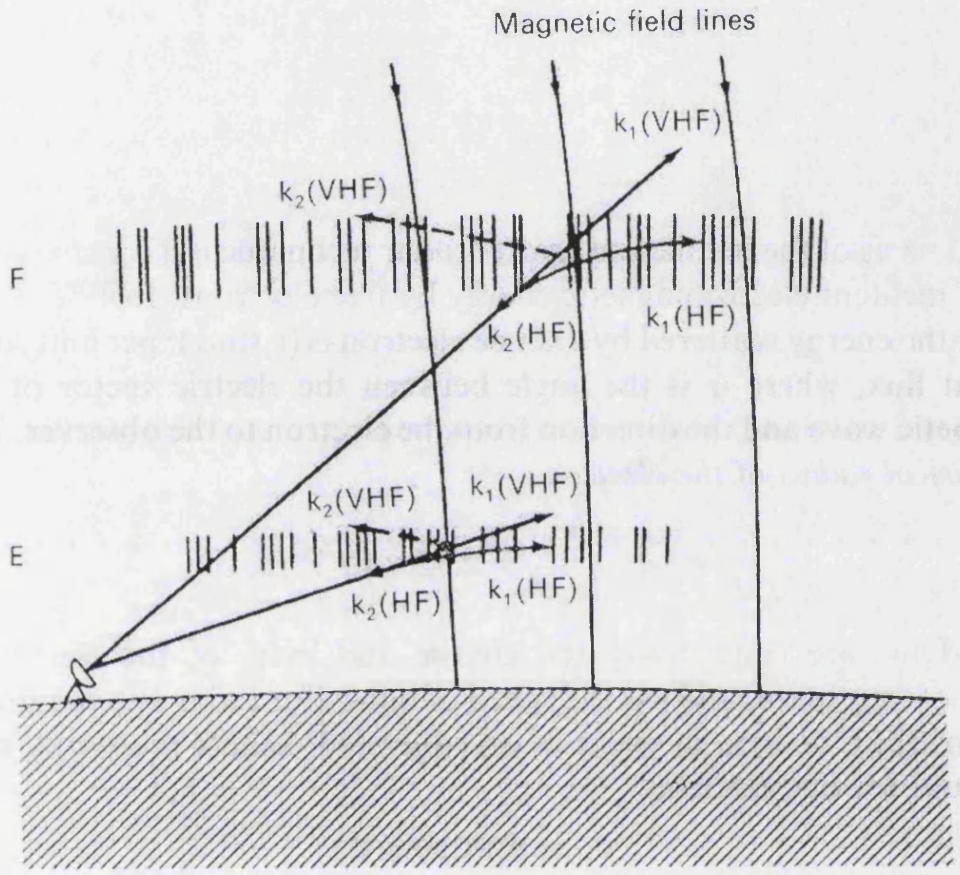


Figure 2.3 Scattering Geometry showing typical paths from the E and F region scatterers of wave vectors k1 and k2 at HF and VHF frequencies (after Greenwald et al., 1985).

In the case of the F-region, HF radars will also reflect off field-aligned irregularities which are due to plasma diffusion along the direction of the field. Consideration of the scattering geometry in the F-region (shown in Figure 2.3) shows that while it is not possible to satisfy the normality condition for the high latitude F region with VHF radar, it is possible to study irregularities by taking into account ionospheric refraction of incident HF radio waves. A problem with this approach however is that the scattering location is not as straightforward to interpret in the F region as for the E region due to the curved ray paths (Hargreaves, 1992).

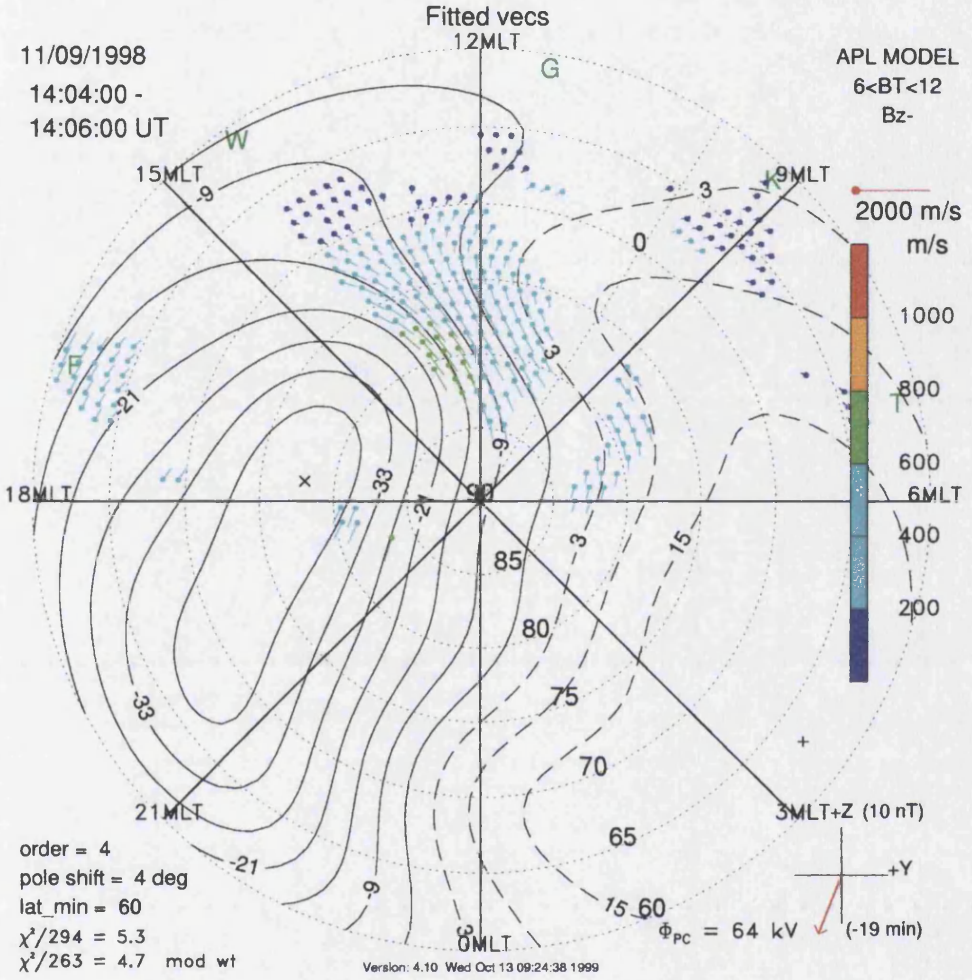


Figure 2.4 Ionospheric convection output from the SuperDARN coherent radar network (courtesy Radio and Space Plasma Physics Group, University of Leicester)

The major parameters measured by the coherent scatter radars are line-of-sight velocity, spectral width and, when used in combination, convection flow vectors. An example of such radars is the CUTLASS system which incorporates radars located in Iceland and Finland. It provides high time-resolution measurements (120 seconds in routine operations) of the ionospheric flow vectors over an area of over 3 million km^2 with a spatial resolution of the order 50 km. A sample of the output possible by using the radar in conjunction with others in the SuperDARN network is shown in Figure

2.4. This Figure shows equipotential contours for an electric field model fit to observed plasma velocities.

2.3.3 Ionosondes and Ionospheric Sounding

The first experimental measurements of the ionosphere were made by Appleton and Barnett in England and Breit and Tuve in the United States in 1924 (Appleton and Barnett, 1925; Breit and Tuve, 1926) following Balfour Stewarts 1882 postulate of the existence of free electrical charges and an electrical current in the upper atmosphere.

The technique of Breit and Tuve has been used since as a standard tool in ionospheric research. The principle of ionospheric sounding is to transmit a radio wave vertically and to measure the time which elapses before an echo is received. This technique has found an application in aircraft detection as radar, which has been an important by-product of early ionospheric physics.

A radio wave propagating through any medium will have its path affected by any free charges. As we have seen in Section 1.3.3 the ionisation in the atmosphere is in the form of several horizontal layers as the electron concentration varies with height. The refractive index is governed by the electron concentration and the magnetic field of the medium and also the frequency and polarisation of the propagating wave. By broadcasting a range of frequencies and measuring the time it takes for each frequency to be reflected, it is possible to estimate the concentration and height of each layer of ionization.

Instruments using these principles are called ionospheric sounders or ionosondes. A pulsed transmitter and receiver are swept synchronously in frequency and the echo delay time is recorded as a function of the radio frequency. The early ionosondes used photographic recording of a cathode ray tube display but since the 1970's it has become usual for instruments to be microcomputer controlled and the data to be recorded and stored digitally.

An idealized daytime ionogram is shown in Figure 2.5 in which only E and F2 layers are present. The existence of the geomagnetic field gives rise to the possibility of two ray paths depending on the sense of polarisation of the transmitted wave. These two ray paths are referred to as the ordinary and extraordinary components.

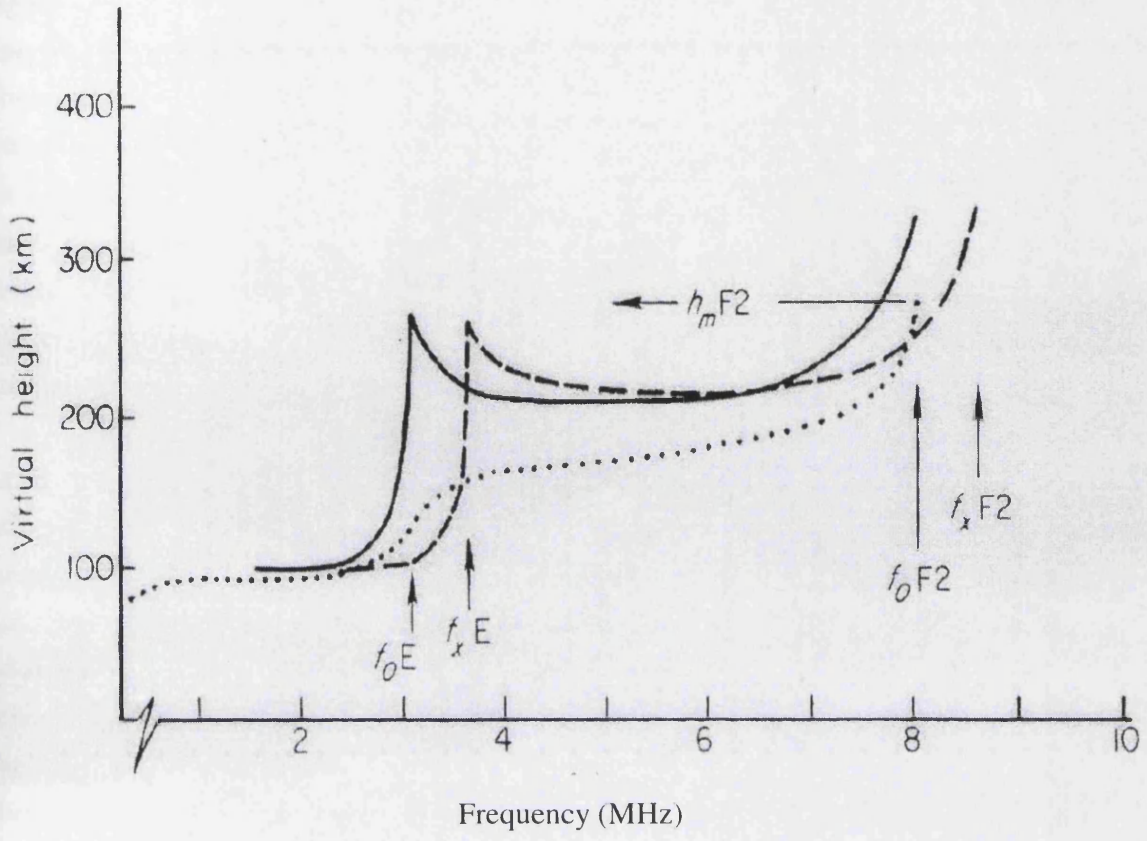


Figure 2.5 An Idealized Ionogram (after Rishbeth and Garriot, 1969).

As the frequency transmitted by the ionosonde increases the waves are refracted less by the ionization in the layer through which they are propagating and so penetrate further before being reflected. As a wave approaches the reflection point the time-of-flight of the signal increases as its group velocity approaches zero. Finally the frequency penetrates the layer and for ordinary mode waves this occurs when the transmitted frequency just exceeds the peak plasma frequency of the layer. In the case of the extraordinary wave, the magnetic field has an additional effect, and reflection occurs at a frequency that is higher than the ordinary wave by half the electron gyrofrequency.

The frequency at which a wave just penetrates a layer of ionization is known as the critical frequency of that layer. In Figure 2.5 the ordinary critical frequency for the E layer is shown, denoted as f_0E with the extraordinary critical frequency denoted f_xE, and similarly there are f_0F2 and f_xF2 frequencies for the F2 layer.

Frequencies transmitted by the ionosonde above this critical frequency will penetrate the layer without being reflected. The ionization encountered passing through the layer will slow the group velocity of these waves however and consequently add to the time-of-flight. Such waves encountering another layer, whose plasma frequency is

higher than the frequency of the wave, will be reflected, and the return signal will be also be delayed on propagation back to the receiver as it passes through the ionization. There will therefore be an apparent, or virtual height indicated by this time delay which will be greater than the true height. The difference between this true height and virtual height is determined by the amount of ionization that the wave has passed through. A procedure, known as true height analysis, can be used to recreate the true-height profile of electron concentration from ionogram data. In Figure 2.5 the dotted curve represents a possible profile of plasma frequency versus real height, which could produce the virtual height curves and the true height is denoted by *hmF2*.

Conventional, analogue ionosondes have made an invaluable contribution to radio science, but they do have limitations based on the fact that only some of the information in the returned signal is used. Some conventional ionosondes do extract amplitude information from the received echoes, but generally only the time delay as a function of radio frequency is measured. Much more information can be derived if the echo phase is measured: e.g. arrival direction, polarization and Doppler shift of the echoes. The development of computer-controlled, digital ionosondes, along with the use of spaced receiving antennae, has made it possible to extract this information and also to scale ionograms automatically without the need for human intervention. Digital techniques have improved the accuracy of frequency measurement by synthesis of signals from stable oscillators and have made data recording and extraction far more efficient and cost effective.

Ionosondes are used extensively for both monitoring the long-term temporal and spatial variations of the ionosphere and for research into space-plasma physics. In a monitoring sounder, the ionograms should be relatively insensitive to the peculiar characteristics of the sounder so that recordings from different sounders in a network can be easily compared. Automatic ionogram scaling and electron-density profile inversion are desirable features in a monitoring sounder as manual ionogram scaling is a highly labour-intensive task and not very practicable when having to scale large amounts of data. A modern digital ionosonde (The Digisonde - Bibl and Reinisch, (1978)) will be described in Chapter 3.

2.4 Models of the Upper Atmosphere

The thermosphere/ionosphere system can be viewed as a highly variable and responsive medium. As such it generally displays a series of classical signatures of the major driving forces of solar UV and EUV heating, and the high latitude geomagnetic forcing resulting from the combination of the magnetospheric electric field and energetic particle precipitation. Changes in the external driving forces, the complex mixture of local, global and non-linear responses of the thermosphere to strong

forcing, the deeply-embedded feedback mechanisms between the thermosphere and the ionosphere, and beyond to the magnetosphere, and the effects of the lower atmospheric disturbances propagating into the thermosphere, combine to cause its observed complex variations. The nature and magnitude of these variations is such that without placing a sequence of observations within the appropriate geographic, solar, seasonal and geomagnetic activity framework, it will be hard to interpret the larger or more exotic variations from an "average" situation.

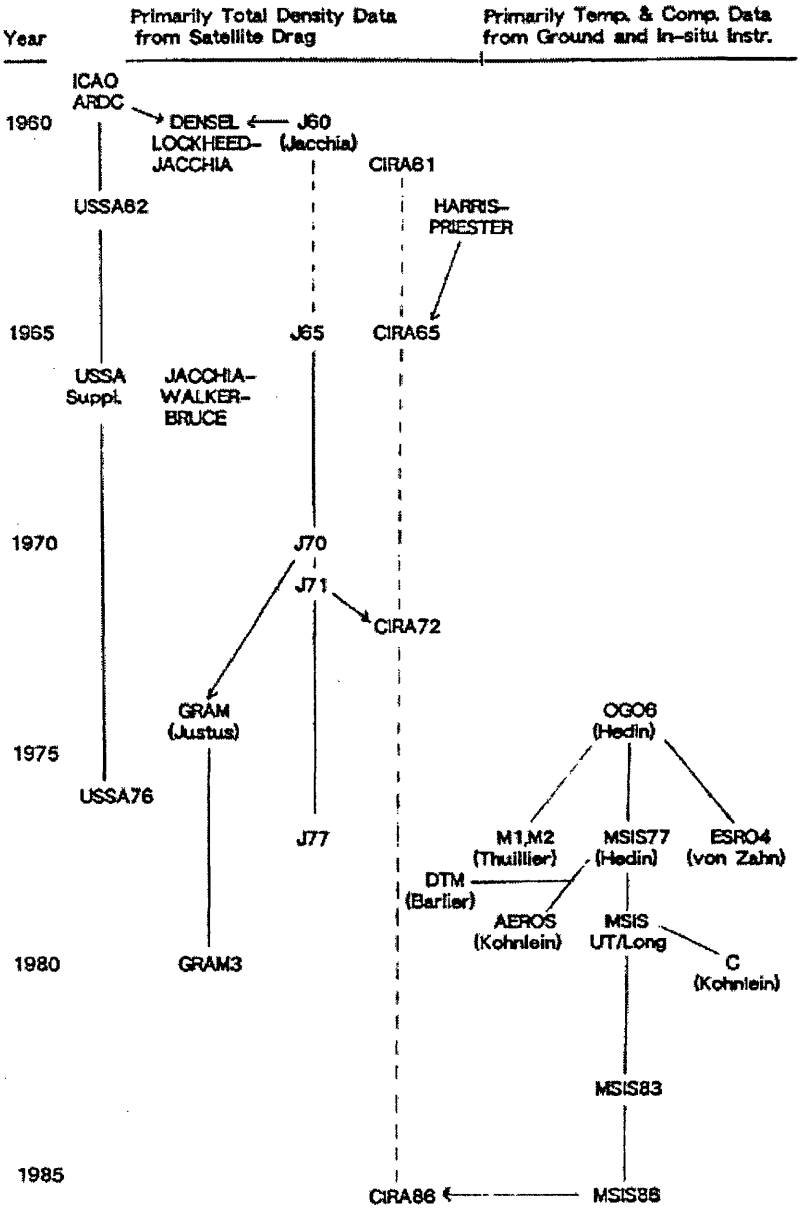


Figure 2.6 Historical development of upper atmosphere models (after Hedin, 1988)

The complexity of both the chemistry and dynamics of the upper atmosphere prevents a complete analytical description of the various parameters, such as

temperatures and wind velocities, to be determined for the entire global upper atmosphere. As the computing power and memory capacity of computers has expanded, models have been developed which attempt to provide descriptions of a limited parameter set for the atmosphere based on empirical observations. The first attempts at developing models were made in the 1960's as satellite drag data became available. L. G. Jacchia developed the first global model of the upper atmosphere by combining satellite data with theory. Figure 2.6 shows the historical development of some of the models up to the Mass Spectrometer - Incoherent Scatter (MSIS-86) model (Hedin, 1987).

The theoretical "harmonic" models of Mayr and Volland (1972), Creekmore et al. (1975), and the two-dimensional models of Richmond and Matsushita (1975), Richmond (1979) and Dickinson et al. (1975) were particularly important milestones, encouraging the development of three-dimensional and time-dependent models (Global Circulation Models or GCM's) as developed by Fuller-Rowell and Rees (1980) and Dickinson et al. (1981).

2.4.1 Empirical Models

Results from the Dynamics Explorer 2 (DE-2) satellite provided new temperature and composition data, and this was used to help produce the MSIS-86 model which included predictions for temperature, density and composition of N, N_2, O, O_2, He, H and Ar with seasonal, daily, annual, magnetic activity, solar activity and longitudinal variations.

Results from the Atmosphere Explorer E and DE-2 satellites were used to generate the HWM87 empirical horizontal wind model (Hedin et al., 1988), analogous to the MSIS models, as the DE-2 satellite provided the first opportunity to obtain wind data over the entire globe. The satellite data consist of zonal winds from the wind and temperature spectrometer (WATS) and meridional winds from the Fabry-Perot interferometer (FPI) on the DE-2 satellite during late 1981 to early 1983, and a mix of meridional and zonal winds from the neutral atmosphere temperature experiment (NATE) on the near-equatorial AE-E satellite during 1975-1978. The NATE data from AE-E range in altitude from 220 to 400 km at low to moderate solar activities. The WATS and FPI data from DE-2 range in altitude from 220 to 600 km at high solar activity. Thus the data included in the model were available for only a limited range of solar activities and altitudes. This led to the model not including solar activity variations, but it provided a basis for future models and was in reasonable agreement with available ground-based measurements for low and mid-latitudes. Miller et al. (1990) showed that that winds inferred from mid-latitude ionosonde data were also in good agreement with HWM87.

A revised model (designated HWM90) that incorporated ground-based data from several incoherent scatter radar and Fabry-Perot optical interferometers, supplemented at the lower boundary by summaries of MF/Meteor radar and older rocket data, extended the data coverage in both solar activity and altitude (Hedin et al., 1991). The satellite data were the same as those used in generating HWM87 (except lower-altitude AE-E data were now used). The basic harmonic expansions used in HWM87 were extended to accommodate the new data and the lower altitude limit of the model was reduced from 220 km to 100 km.

Generally, the Fabry-Perot measurements included in the model are taken only during local night and assume the vertical wind is zero. Incoherent scatter F region data were available for all local times, but nighttime data were expected to be less accurate than daytime data because of corrections for ambipolar diffusion, based partly on quantities estimated from models, which are more important at night.

Both MSIS-86 and HWM-90 have been revised and extended to include mesosphere and stratosphere coverage in MSIS-90 and HWM-93. No significant changes were made to the treatment of the upper atmosphere however, indicating the lack of any larger satellite datasets becoming available, leading to MSIS-86 and HWM-90 still being used as reference empirical models by upper atmosphere researchers (e.g. Dyson et al., 1997).

The International Reference Ionosphere (IRI) is an international project sponsored by the Committee on Space Research (COSPAR) and the International Union of Radio Science (URSI). These organizations formed a working group in the late 1960's to produce an empirical standard model of the ionosphere, based on all available data sources. Several steadily improved editions have been released. For given location, time and date, IRI describes the electron density, electron temperature, ion temperature, and ion composition in the altitude range from about 50 km to about 2000 km. The major data sources are the worldwide network of ionosondes, incoherent scatter radars, topside ionosphere sounders and in situ instruments on several satellites and rockets.

2.4.2 General Circulation Models

Development of 3-dimensional time-dependent models became possible in the early 1980's, and has led to two models which are sufficiently developed for use in analysing data and for the interpretation of physical processes responsible for the observed thermospheric structure and dynamics. The United States National Center for Atmospheric Research Thermospheric Global Circulation Model (NCAR-TGCM) (Dickinson et al., 1981) and University College London Thermospheric Global Circulation Model (UCL-TGCM) (Fuller-Rowell and Rees, 1980) both solved the

primitive equations of dynamic meteorology, modified for thermospheric dynamics, and include coupling between composition and dynamics. This provided a three-dimensional, time-dependent model of the Earth's neutral upper atmosphere, using a finite differencing technique to obtain a self-consistent solution for the coupled, nonlinear equations of hydrodynamics, thermodynamics and continuity of the neutral gas and for the coupling between the dynamics and the composition. The NCAR-TGCM model uses a grid with an effective separation of 5 degrees in both latitude and longitude. In the vertical plane there are 24 constant-pressure surfaces located at heights between 97 and 500 km, a spacing corresponding to half the scale heights through this altitude range. Model predictions have been checked against experimental data from the DE-2 satellite (Roble et al., 1988a) and also against ground-based measurements (Meriwether et al., 1988) both of which produced good agreement. To this extent the numerical procedures used in the NCAR-TGCM and the UCL-TGCM may be considered to have been validated using comprehensive global-scale experimental data sets (Burns et al., 1991).

Early studies with three dimensional models (Fuller-Rowell and Rees, 1980, 1981; Rees et al., 1980; Roble et al., 1982) showed that the dynamical response of the thermosphere to the energy and momentum sources associated with geomagnetic forcing could be crudely simulated by including simplified models of the polar convection electric field such as those of Heppner (1977), as parameterized by Volland (1979), and using the global Chiu (1975) ionospheric model.

The first attempt at computing a self-consistent ionosphere and thermosphere model (Quegan et al., 1982; Fuller-Rowell et al., 1984) was steady-state, with no time dependence, and was derived for northern polar winter by iterating between the UCL-TGCM and a separate ionospheric code developed at Sheffield University. This iterative approach proved valuable, the data from the polar ionosphere model being used in several thermospheric model simulations. Subsequently, fully coupled, self-consistent models of the ionosphere-thermosphere system were developed for the entire globe (Roble et al., 1988b; Fuller-Rowell et al., 1988) when the UCL-TGCM became CTIM (Coupled Thermosphere Ionosphere Model), with the NCAR-TIGCM (I for Ionospheric) being further developed by inclusion of electrodynamics into the NCAR-TIEGCM (Richmond et al., 1992). Comparisons with data and with empirical models show that the major diurnal, latitudinal and seasonal variations of the mid- and low latitude thermosphere, in wind velocity, temperature and density of the major species, can be successfully represented in theoretical models for a wide range of solar and geomagnetic activity levels.

A spectral representation of NCAR-TIGCM simulations has been used to produce the Vector Spherical Harmonic (VSH) model (Killeen et al., 1987) which

allows global thermospheric and ionospheric parameters to be derived. While this allows parameters such as neutral wind and temperature to be derived without the much higher computational overhead of a full NCAR-TIGCM run being carried out, the VSH model results are a series of interpolations of simulations of the wind and temperature field from the NCAR-TIGCM for given stationary conditions rather than a fully time-dependent prediction. This can lead to disagreements in comparisons of specific experimental data to VSH output (Smith et al., 1994).

2.4.3 The Field Line Interhemispheric Plasma (FLIP) Model

The one-dimensional FLIP model (Richards, 1991) has been developed since the early 1980's, its main component calculating the plasma densities and temperatures along entire magnetic flux tubes from 80 km in the northern hemisphere through the plasmasphere to 80 km in the southern hemisphere. The model uses a tilted dipole approximation to the Earth's magnetic field and the continuity and momentum equations for O^+, H^+ and He^+ are solved, as formulated for the topside ionosphere by St. Maurice and Schunk (1977). Including collisions between ions and neutrals allowed the equations to be extended into the E and F regions.

Within the model the energy equations of Schunk and Nagy (1978) are solved to obtain the electron and ion temperatures and the method of Nagy and Banks (1970) is used to solve the two-stream photoelectron flux equations, thereby providing the electron heating due to photoelectrons. Richards and Torr (1988) used updated cross sections and solar EUV fluxes as inputs to demonstrate that the model photoelectron fluxes were in good agreement with the measured fluxes of Lee et al. (1980).

2.5 Ion-Neutral Coupling in the Thermosphere

The strong coupling of the thermospheric neutral and ionospheric plasma environments has been well established. Rishbeth and Setty (1961) noted that ionospheric F-region plasma densities at mid-latitudes in the winter were higher than in summer, and were the first to suggest that this "winter anomaly" was due to a change in the ratio of the neutral atomic oxygen to molecular nitrogen. Foster (1984) has presented average summer and winter F-region densities from the Chatanika incoherent scatter radar, which indicates a high latitude (60^o - 75^o) "winter anomaly", confirmed by results from the EISCAT radar indicating a similar seasonal behaviour at European longitudes (Farmer, 1987).

The high latitude neutral thermosphere is strongly coupled to the ionosphere through collisions between neutral and ionized species. At these latitudes, the direct electrical coupling of the magnetosphere with the ionosphere induces F-region ion convection, which is strongly impressed upon the neutral gas through collisions. Thus

solar wind effects on the magnetosphere, such as the effects of varying interplanetary magnetic field (IMF) orientation, are subsequently manifested in the neutral thermosphere circulation through this relationship. In the F-region, resonant interaction between neutral and ionized atomic oxygen is the dominant collisional process. Since the O^+-O collision frequency is proportional to the electron density, the altitude of maximum ion-neutral coupling in the thermosphere will generally occur at or near the F-region electron density peak.

Geomagnetic control of the ionospheric and neutral thermospheric constituents in the polar regions has been demonstrated in the observations from the DE-2 satellite reported by Hays et al. (1984) and Killeen et al. (1983).

2.5.1 Rishbeth Servo Model

The behaviour of the F2-layer under the influence of production and loss, diffusion and drift has been studied by numerous authors. From earlier work (e.g. Duncan, 1956) there developed the idea that the F2-layer behaves like a servo system obeying the following laws (Rishbeth, 1967):

(A) In the absence of applied vertical drift, the F2 peak lies at a "balance height" determined by diffusion and loss.

(B) Vertical drift due to electric fields or neutral winds displaces the equilibrium position of the peak to a new level, which is time-varying if the drift is time-varying.

(C) At any instant the actual height of the peak approaches its equilibrium value at a rate determined by diffusion and loss.

(D) The rate of change of peak electron density is determined by local values of the production rate and loss coefficient.

Following the treatment of Miller et al. (1997) a servo equation that describes the steady state meridional wind speed as a function of F2-layer peak height (*hmF2*) can be written:

$$U = \frac{D \sin I}{2H \cos I}\left(e^{-k[(hmF2-h_0)/H]} - e^{(hmF2-h_0)/H}\right) \qquad (2.3)$$

where U is the poleward horizontal component of the neutral wind along a geomagnetic meridian and I is the magnetic dip angle. The scale height factor k is approximately equal to 1.875, the ratio of the scale height of O to the average scale height of O_2 and N_2 (Rishbeth, 1967; Buonsanto et al., 1989). This scale height factor has the effect of changing the scale height, H, from the atomic oxygen scale height that is appropriate for the diffusion term to the molecular scale height, appropriate for the

recombination term. The diffusion coefficient, D, is determined at the balance height h_O, the height of the F2-layer peak in the absence of a vertical drift due to meridional wind or zonal electric field. This balance height needs to be computed to be compared to a measured *hmF2* and hence calculate the equivalent wind.

Using this technique does however introduce assumptions that the F region is horizontally stratified and that vertical motions are small, which limit its application to locations and conditions which fulfil these criteria. Upwelling due to auroral activity at high latitudes would for example produce strong vertical flows that would in turn affect the height of the F layer as well as the computed winds. The equatorial fountain effect at low latitudes would produce similar effects so there is a need to consider the conditions and location carefully before applying the servo theory to derive neutral winds.

2.5.2 Meridional Wind Model

The possibility of using the long term database of ionosonde mesurements, available from numerous stations around the world, to derive meridional neutral winds led to speculation (e.g. Buonsanto, 1990) that it might be possible to provide an empirical model of these winds based on the ionosonde database. The Meridional Wind Model (MWM), of Miller et al. (1997), described in this section was produced in part to provide a near-global validation of the General Circulation Model meridional wind output and also to provide a convenient method of wind estimation for other models. With the long-term ionosonde data base, it also provides the capability of making extended climatological studies of the ionosphere. For any predictive model of the ionosphere the meridional wind component must be computed or be otherwise available otherwise any modelled ionospheric profile will be inaccurate (Miller et al., 1997).

The MWM method of deriving winds from the F2 region peak density (*hmF2*) combines the basic concepts of the servo model with the more sophisticated photochemistry of a thermospheric model. The FLIP ionospheric model (Richards, 1991) is used to determine the relationship between *hmF2* and the meridional neutral wind within the MWM. This relationship is determined as a function of time for the appropriate day, location, and solar conditions, and is used to convert the difference between *hmF2* and the balance height to an equivalent wind speed as was introduced in the previous section.

In order to take advantage of the treatment of the photochemistry contained within the FLIP model, Miller et al. (1997) simplify Equation (2.3) by recognizing that the two exponentials in Equation (2.3) are of comparable magnitude only near the balance height. If the measured value of *hmF2* is more than a scale height above or

below the balance height, then the first or second term dominates in the equation and the remaining term becomes unimportant. As a result in the following example, for an equatorward wind and a layer height that is greater than the balance height, Equation (2.3) becomes

$$U = Ce^{(hmF2-h_0)/H} \tag{2.4}$$

where C is constant at a specific location and at specific solar and magnetic activity levels. This constant may be determined by using the FLIP model to specify an equatorward wind speed, U_q, and to find the corresponding layer height, h_q, such that

$$C = U_q e^{-(h_q-h_0)/H} \tag{2.5}$$

Using this result Equation (2.4) can then be written

$$U = U_q e^{(hmF2-h_q)/H} \tag{2.6}$$

Using the FLIP model to specify a poleward wind, U_p, a similar expression with a corresponding layer height, h_p, can be derived. Using these two terms together gives an expression that closely resembles the servo equation:

$$U \cong U_p e^{-k[(hmF2-h_p)/H]} + U_q e^{(hmF2-h_q)/H} \tag{2.7}$$

Miller et al. (1993) showed that, in spite of the approximations employed, Equation (2.7) reproduced very accurately the height of the F layer height at a given meridional wind speed as calculated using the photochemical equations of the FLIP model. Miller et al. (1997) point out that the advantage of Equation (2.7) is that h_p and h_q can be calculated separately using the FLIP model for specified U_p and U_q, and then substituted into Equation (2.7). In this way the full potential of the photochemistry of the model is available to the MWM, without the uncertainty of the approximations needed for the calculation of balance height and the dawn and dusk transitions of the servo equation. Further, Equation (2.7) makes a smooth transition at small wind speeds where neither exponential is dominant and, where U is small, Equation (2.7) reduces to a linear equation.

Three sources for heights of *hmF2* are allowed in the MWM. A time series of *hmF2* values may be input directly from an ASCII table of measurements, the IRI model may be run to derive values of *hmF2* or ionosonde parameters may be used

from which the *hmF2* value can be derived by means of the method of Dudeney (1983).

This latter method is included as an option because *hmF2* is not a routinely recorded parameter from ionosonde data, due to the labour-intensive nature of such a derivation. Only a small fraction of all ionograms recorded since the 1930's have ever been analysed to derive *hmF2* values using true height analysis.

It is possible however to make use of the inverse relationship between the routinely scaled *M(3000)F2* parameter and the *hmF2* height to use ionosonde parameters in the MWM. *M(3000)F2* is defined as the ratio of the maximum usable frequency and the F2-layer ordinary mode critical frequency (*foF2*). Dudeney (1983) published an expression developed from earlier work by Shimazaki (1955) and Bradley and Dudeney (1973) which is used in the MWM to relate the *M(3000)F2* parameter to a value of *hmF2*.

In this manner ionosonde data stored in the International Union of Radio Science Ionospheric Informatics Working Group (IIWG) format for ionosonde parameters may be used as input to the MWM. As mentioned above, if neither source of experimental data is available the IRI may be run as a sub-routine of the MWM to derive *hmF2* values.

2.6 Limits of Current Theoretical Predictions:

A number of problems exist in trying to model ion-neutral coupling and also in applying model predictions to specific campaigns and experiments. Two of the more fundamental problems are the uncertainty in the value of the ion-neutral collision cross-section and the poor performance of models in representing the *O* density at thermospheric altitudes.

Atomic oxygen dominates the composition of the earth's thermosphere above about 200 km, and as a result of the Sun's ionizing radiation, atomic oxygen ions dominate the composition of the earth's ionosphere. Thus, interactions between O^+ ions and their parent gas control the structure and dynamics of the earth's upper atmosphere, and proper representation of parameters such as the O^+-O collision frequency for momentum transfer and charge exchange becomes crucially important in models of the upper atmosphere and in the derivation of atmospheric results from experimental measurements.

The impact of uncertainties in the O^+-O momentum transfer collision frequency on the behaviour of the ionospheric F-region has been studied by Moffett et al. (1990). A 70% increase in the collision frequency was found to increase the peak electron density in the F-region by as much as 25% at night and by 10% in the daytime, as well as raise the height of the maximum density by 20 km at night. The influence of

the O^+-O collision frequency on the structure of the neutral upper atmosphere has been studied parametrically by Roble (1988). Increases of 70% in the atomic oxygen collision frequency were found to increase the exospheric neutral temperature by as much as 80 K in the daytime, and to increase globally averaged Joule heating by about 80%. Because frictional drag between ion and neutral particles is proportional to the collision frequency, it is found that the global neutral wind pattern is also altered, with changes in the wind speed as large as 100 ms^{-1} at middle and high latitudes.

Techniques have been developed and applied to determine the meridional neutral wind component in the thermosphere using incoherent scatter radar measurements of ionospheric drifts (e.g. Salah and Holt, 1974), and using ionosonde measurements of the height of the peak ionospheric F-layer density (e.g. Miller et al., 1986, 1997).

The derivation of the neutral wind based on ionospheric measurements depends strongly on the knowledge of the ion diffusion velocity which, in turn, is dependent on the ion-neutral collision frequency for atomic oxygen. The theoretical work of Roble (1988), referred to above, was confirmed in the empirical calculations of Buonsanto et al. (1989) where neutral winds derived from the Millstone Hill ISR measurements had differences as large as 105 ms^{-1} at local midnight when the atomic oxygen collision frequency was increased by 70%.

As a result of the experimental evidence to hand (Burnside et al., 1987; Sipler et al., 1991; Buonsanto et al., 1992; Davis, 1993), an interim standard for the O^+-O collision frequency was adopted of 1.7 times the early values derived by Dalgarno (1964), Banks (1966) and Schunk and Walker (1973) which were primarily based on laboratory measurements. This factor of 1.7 has been dubbed the "Burnside factor" and was until recently a consensus standard for the Coupling, Energetics and Dynamics of Atmospheric Regions (CEDAR) community but also widely used in the aeronomical community at large. Omidvar et al. (1998) have now suggested that the factor is in fact 0.98 with +/- 8 % uncertainty and conclude the factor of 1.7 was due to a 30 % underestimation of the calculated collision frequency by Dalgarno (1964) and a 40 % over-estimation of the empirical determination by Burnside et al. (1987) due to accumulative statistical errors.

Unfortunately neither of the two global wind models, HWM90 or VSH, is yet capable of providing meridional neutral winds with sufficient precision for accurate ionospheric modelling (Richards, 1991). The HWM90 model suffers from a lack of data, especially in the southern hemisphere, while the VSH model suffers from the imprecise knowledge of the inputs to the NCAR-TIGCM model. Even if the wind models were adequate on an average basis, they would be unlikely to be precise enough for ionospheric modelling of specific data sets owing to the highly variable

nature of the upper atmosphere. In addition, ionospheric modellers would still need information on the electric fields which is also difficult to obtain.

2.7 Combined Measurements of Ion-Neutral Coupling

In attempting to investigate ion-neutral coupling it is inevitable that combined measurements from instruments determining ionospheric and neutral parameters be used. While it is possible for incoherent scatter radars to estimate the neutral wind, these measurements are indirect and depend on the value used for O^+ diffusion velocity which in turn is dependent on the value of the O^+-O collision frequency. As this is currently subject to considerable speculation the derived neutral winds must be regarded as less accurate than direct measurements from FPIs. The same problems apply to derivations of the neutral wind from ionosonde data. Experimenters have sought to devise combinations of instruments to allow comparison of the measurements from these instruments to establish a better understanding of the limits of our current theoretical framework for ion-neutral coupling.

2.7.1 Ionosonde - FPI Comparisons

The first reported comparison of FPI and ionosonde data using servo theory principles was for the southern hemisphere mid-latitude by Yagi and Dyson (1985). They found reasonable overall agreement between observed changes in the height of the nighttime F-layer and the changes calculated from servo theory using neutral winds from Fabry-Perot measurements.

Full diurnal calculations of the horizontal wind from ionospheric electron density profiles were first given by Miller et al. (1986, 1987) using values of *hmF2* obtained from the true height analysis of ionograms at Arecibo (30^oN geomagnetic), where the results were compared with meridional winds from Fabry-Perot interferometer and incoherent scatter radar data, revealing good agreement.

Gurubaran and Sridharan (1993) used the servo equations to estimate the balance height of the F-region ionization maximum by inputting neutral wind and temperature data from an FPI and comparing these with independently measured *hmF2* by means of ground-based ionosonde, at Mount Abu/Ahmedabad, India, a low latitude site. The comparison (e.g. Figure 2.7) reveals fairly good agreement, reproducing most of the observed features and providing direct experimental evidence for the two regions to behave as a closely coupled system. They note that the influence of electric fields, though they appear to be less significant at this latitude, could still be inferred from the data.

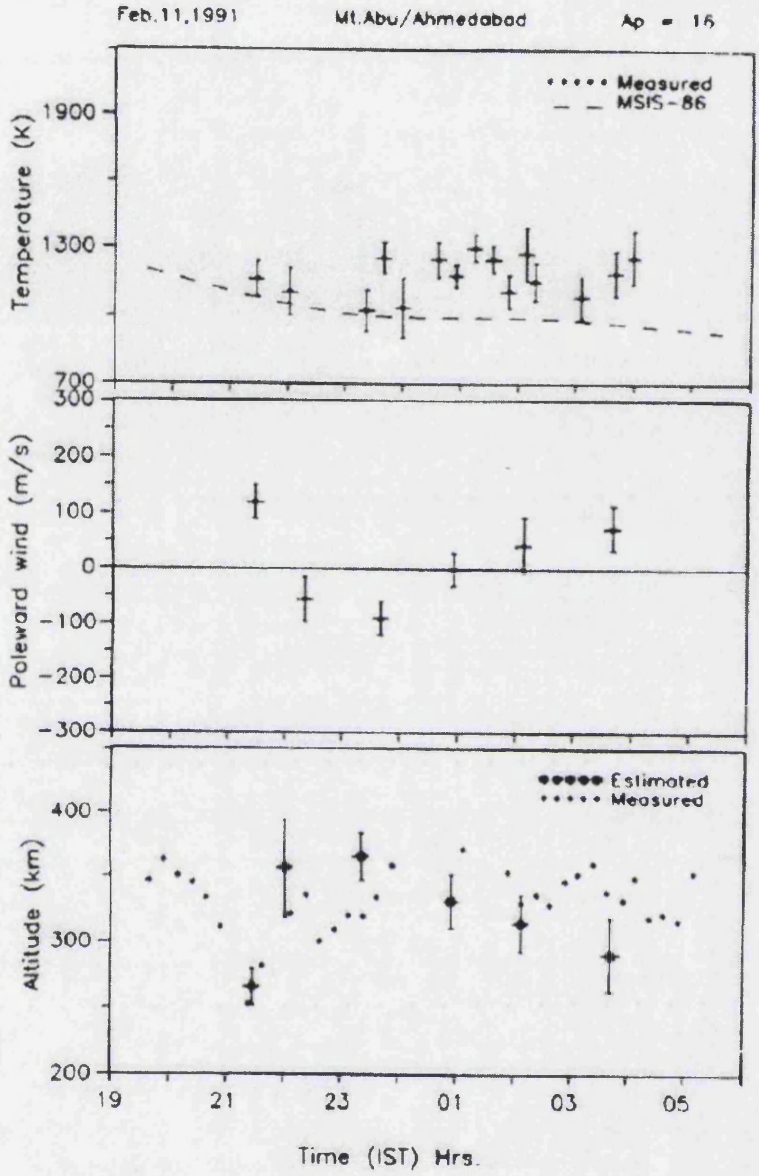


Figure 2.7 FPI - Ionosonde comparisons at low latitude (after Gurubaran and Sridharan, 1993)

Thayer et al. (1995) combined the use of an FPI at Thule Air Base, Greenland with a digital ionosonde located at Qanaq, Greenland to obtain a continuous record of F-region neutral winds, electron density profiles and F region ion drifts at high latitude. This combination of ground-based observations allowed the investigation of ion/neutral coupling at a temporal resolution of about 15 min. They investigated the observed response of the neutral wind to convection changes in the ion drift inside the polar cap for southward and northward IMF B_Z conditions, and concluded that the ion drag term alone could not describe the observed response of the neutral wind during northward IMF.

Recently Dyson et al. (1997) have reported results at southern mid-latitude from comparison of FPI neutral winds with those derived from Digisonde measurements of *hmF2*. They find that in implementing the servo equations the best fits between the data sets are found by using a Burnside factor of 2.0, as opposed to the recommended factor of 1.7 . They suggest that lower values of the Burnside factor could only be tolerated by inferring large electric fields or large discrepancies in the atomic oxygen densities derived from the MSIS-86 model.

Previously no direct comparisons of FPI neutral winds to neutral winds derived from ionosonde data have been performed at high latitudes. This is probably due to the difficulty in taking possibly large electric fields into account. An experimental arrangement to perform this comparison at Kiruna/Tromso is described in the following chapter.

2.7.2 EISCAT INDI Experiment

A further example of the experimental possibilities for investigating ion-neutral coupling is the EISCAT Ion Neutral Dynamics Investigation (INDI) experiment. The EISCAT radar system will be described in detail in Chapter 3. This experiment resulted from early attempts to compare FPI and EISCAT ISR data being complicated by the lack of coordination between the two instruments (Rees et al., 1984). For the INDI experiment an FPI is used at Kiruna. Early versions of the INDI experiment attempted to scan the radar through 360^{o} every 15 min, observing the same volumes as the FPI. This was not possible as the radar had to return through 360^{o} after each scan, which took too long.

As a result, it was decided that the radar need scan only through the meridian, which allows the determination of the meridional component of the neutral wind, the main object of the comparative study. A modified version of the EISCAT CP-3 experiment was used (Rishbeth and Williams, 1985), with only the seven central positions of the original scan from the north of Tromso to above Kiruna. In this way, the scan time was reduced to 15 min from a standard CP-3 scan time of 30 min, each of the seven measurements taking approximately 2 min. Figure 2.8 is a plan of the INDI experiment geometry. This shows the relation between the FPI and EISCAT scattering volumes.

The results of the experiment have been used to estimate the O^{+}-O collision parameter by adjusting the value used in the derivation of winds from the radar until the best fit was achieved. It has been pointed out that much care is necessary in taking into account the variability of the data used and some re-evaluation of the methods used has been undertaken in similar experiments as a consequence (Davis et al., 1995; Buonsanto et al., 1997c). The values of the Burnside factor derived using this

technique tend to be around 1.2, significantly lower than the recommended value of 1.7 (Salah, 1993) and in keeping with the recent trend to downward revision of the Burnside factor.

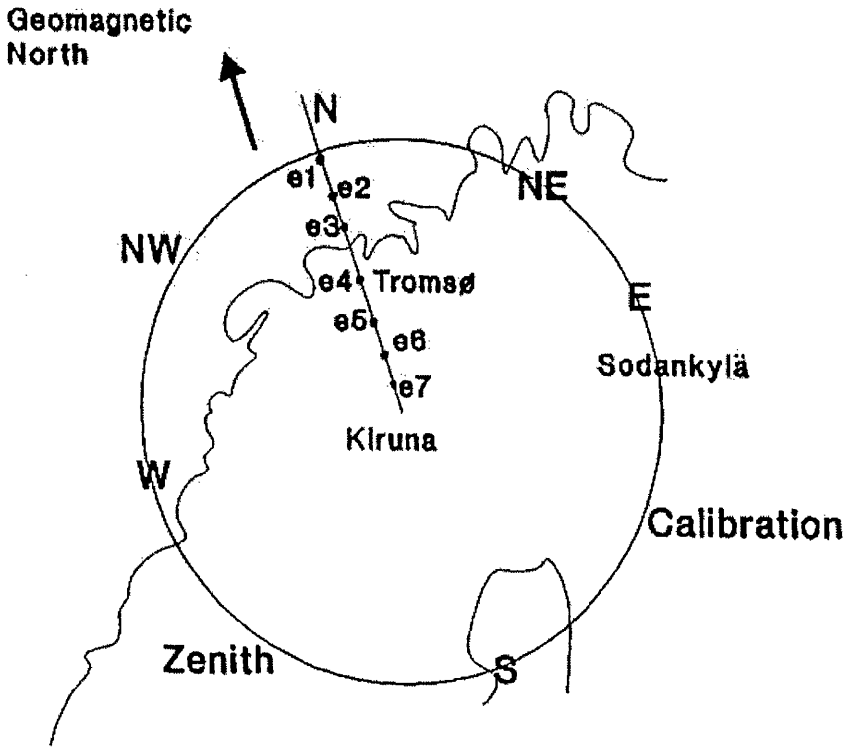


Figure 2.8 Plan view of experimental geometry (after Davis et al., 1995)

2.8 Geomagnetic activity indices

Several indices of geomagnetic activity are used to classify periods according to their level of activity. The most often used are A_p and K_p, both based on magnetometer data. K_p is based on the range of variation within 3-hour periods of the day (0000-0300 UT, 0300-0600 UT, etc.) observed in the records from about a dozen selected magnetic observatories. After local weighting, and averaging, the K_p value for each 3 hours of the day is obtained on a scale from 0 (for very quiet) to 9 (for very disturbed). The scale is quasi-logarithmic, and the integer values are subdivided into thirds by use of the symbols + and -: thus, 3, 3+, 4-, 4, etc.

A_p is a daily index, obtained from the same basic data, but converted to a linear scale (the 3-hour a_p) and then averaged over the day (0000-2400 UT). The value of the intermediate a_p is approximately half the range variation of the most disturbed magnetic component, measured in nanoteslas. The relationship between K_p and a_p is as in Table 2.1

Table 2.1 Relationship between K_p and a_p.

K_p	a_p
0	0
1	3
2	7
3	15
4	27
5	48
6	80
7	140
8	240
9	400

2.9 Solar activity indices

Similar to the A_p and K_p indices for geomagnetic activity there are indices available to gauge solar activity, the most often used in upper atmosphere physics are the $F_{10.7}$ index and the sunspot number (SSN).

Solar radio flux, which originates from atmospheric layers high in the sun's chromosphere and low in its corona, changes gradually from day to day in response to the number of spot groups on the disk. Solar flux from the entire solar disk at a frequency of 2800 MHz (10.7 cm) has been recorded routinely by radio telescope near Ottawa since February 1947. The observed values are adjusted for the changing Sun-Earth distance and for uncertainties in antenna gain giving an $F_{10.7}$ index in units of 10^{-22} J s^{-1} m^{-2} Hz^{-1}.

Yearly sunspot numbers are available since the telescope was invented in 1610. The sunspot number index (R) is a measure of the area of solar surface covered by spots. As the number of spots increases and their magnetic complexity grows, they become likely sources of large eruptive energy releases known as solar flares. Besides daily, monthly, and yearly averages, a 12-month running mean value (R12) is calculated. In 1842, H. Schwalbe noticed that the number of spots varies with a period of about 11 years. Analyzing the long record of observations, it has been found that the "solar cycle" period varies from ten to 12 years with a mean of 11 years.

2.10 Summary

Many techniques have been developed to measure the most important neutral and ionospheric parameters in the upper atmosphere. These are of great importance in allowing long-term trends to be established and also for direct measurements of specific phenomena in campaign based programmes. This importance is enhanced by the lack of satellite coverage in general for these important parameters at present.

It is now possible to simulate the time-dependent response of the terrestrial thermosphere and ionosphere, and their interactions, under combined solar UV/EUV and geomagnetic forcing, as described by theoretical and semi-empirical models, in self-consistent numerical models. The numerical models are in basic agreement with observations, and with semi-empirical thermosphere models, and can thus be used to analyse the major thermospheric driving forces. These models provide a means of investigating the behaviour of the thermosphere and ionosphere under its normal range of activity, and also a way of attempting to understand what may have happened prior to some of the more unusual events which have been reported.

The development of an accurate neutral wind climatology for the F-region would be an important asset in studies of other ionospheric characteristics. A wind climatology is difficult to obtain because of the difficulty in making global measurements at F-region altitudes. The Meridional Wind Model described in this chapter provides a means of obtaining an estimate of the component of the neutral wind parallel to the magnetic meridian. Because of the strong influence of the magnetic field on the ionosphere at F-region heights, the meridional component of the wind must be known for other ionospheric characteristics to be modelled. This capability, in turn, is essential to the development of models to provide accurate forecasts of ionospheric conditions.

A recent intercomparison of the performance of five physical models of the upper atmosphere has been published (Anderson et al., 1998) which includes three of the models described here, the NCAR-TIGCM, CTIM and FLIP. The original motivation for the study was to determine why several of the models consistently underestimate the F region peak electron density in the mid-latitude daytime ionosphere at solar maximum. The resolution of this problem did not reveal any lack of physics in the models but instead the need for adjustment in various chemical reaction rates, photoionization processes, and diffusion coefficients, with the main one being the adoption of the Burnside factor of 1.7 for the diffusion coefficients.

Specific experiments have been devised and performed to help to address the outstanding problems in our understanding of the coupling of the ionosphere and thermosphere, with considerable interest in establishing a reliable value for the Burnside factor, as it is vital in generating accurate results for any such investigation.

The development of the MWM allows relatively easy comparison between ionosonde and FPI datasets and should allow meridional neutral winds to be derived from the more numerous ionosonde sites around the world. However high latitude investigations using the MWM are complicated by the FLIP model restrictions on applicability, and the influence of electric fields which are not accounted for by the model.

Chapter 3 - Experimental Setup at Mid-latitude and High latitude Sites

3.1 Introduction

This chapter will describe the detail of the instrumentation and configuration used to provide the data for this thesis. As the Fabry-Perot Interferometer (FPI) data has been produced by the Atmospheric Physics Laboratory (APL) at University College London (UCL) the use of this type of instrument will be described in detail in relation to thermospheric physics.

3.2 Temperature and Wind Measurements using a Fabry-Perot Interferometer

A theoretical treatment of the FPI is given in Appendix A and its characteristics described. The high resolving power of the FPI makes it an extremely useful instrument to use in the study of the optical atmospheric emissions discussed in Chapter 1. In 1923 Babcock used the FPI in its first geophysical application by making the first wavelength determination of the green line of the aurora, now known to be the 557.7 nm transition of atomic oxygen. In using the FPI in a spectroscopic mode it is possible to determine both the temperature of the emitting species and the component of wind velocity along the line of sight of the observations.

An emission line profile will be broadened as a result of the Doppler effect, where the apparent wavelength of the emission from an atom or molecule will be shifted if it has a component of velocity along the line-of-sight of the observer. In making observations of the 630 nm emissions from atomic oxygen, the line profiles will be broadened by the random thermal motions of the emitting atoms. A Doppler shift in the wavelength emitted by each atom is given by

$$\frac{-\Delta\lambda}{\lambda_0} = \frac{\Delta\upsilon}{\upsilon_0} = \frac{v_x}{c} \qquad (3.1)$$

where v_x is the velocity of the emitting atom toward the observer, c is the speed of light and $\Delta\upsilon$ is the frequency shift. The kinetic theory of gases expresses the temperature of a gas in terms of the random motion of the constituent molecules (or atoms for an atomic species) as follows:

$$\frac{1}{2}m\overline{v^2} = \frac{3}{2}kT \tag{3.2}$$

where m is the mass of one atom, T is the absolute temperature and $\overline{v^2}$ is the mean square velocity of the assembly of atoms. The atoms within a body of gas exchange energy by collisions and reach a state of thermal equilibrium, resulting in the velocities being distributed according to the Maxwell-Boltzmann law:

$$\frac{dN(v_x)}{N} = \frac{\alpha}{\sqrt{\pi}}\exp\left(\frac{-v_x^2}{\alpha^2}\right)dv_x \tag{3.3}$$

where $\frac{dN(v_x)}{N}$ is the fraction of atoms with velocity between v_x and $v_x + d\,v_x$, α is the most probable velocity given by

$$\alpha = \left(\frac{2RT}{M}\right)^{\frac{1}{2}} = \left(\frac{2kT}{m}\right)^{\frac{1}{2}} \tag{3.4}$$

where M is the mass number for the species and R is the universal gas constant. By substitution from Equation (3.1) into Equation (3.3) it is found that the fraction of atoms emitting in the frequency range υ to $\upsilon + \mathrm{d}\,\upsilon$ will be given by

$$\frac{dN(\upsilon)}{N} = \frac{1}{\alpha\sqrt{\pi}}\exp\left(\frac{-c^2(\Delta\upsilon)^2}{\upsilon_0^2\alpha^2}\right)\frac{c}{\upsilon_0}d\upsilon \tag{3.5}$$

Noting that the intensity is proportional to $dN(\upsilon)$, the full range of frequencies for the line profile can then be expressed as follows,

$$I_\upsilon = I_0\exp\left(\frac{-c^2(\upsilon_0-\upsilon)^2}{\upsilon_0^2\alpha^2}\right) \tag{3.6}$$

where I_0 is the central intensity (Thorne, 1988). The values of frequency at which the intensity is half of that at the centre, denoted $\upsilon_{½}$, will have $I_\upsilon = I_0/2$. Substituting this value of intensity into Equation (3.6) and taking the logarithm of both sides

$$\frac{-c^2\left(\upsilon_0-\upsilon_{\frac{1}{2}}\right)^2}{\upsilon_0^2\alpha^2}=\ln\left(\frac{1}{2}\right) \tag{3.7}$$

which leads to the full width at half maximum (FWHM) being

$$\delta\upsilon_D=\frac{2\upsilon_0}{c}\sqrt{\left(\frac{2RT\ln 2}{M}\right)}=7.16\times10^{-7}\sqrt{\left(\frac{T}{M}\right)}\upsilon_0 \tag{3.8}$$

and in dimensionless form

$$\frac{\delta\upsilon_D}{\upsilon_0}=\frac{\delta\lambda_D}{\lambda_0}=7.16\times10^{-7}\sqrt{\left(\frac{T}{M}\right)} \tag{3.9}$$

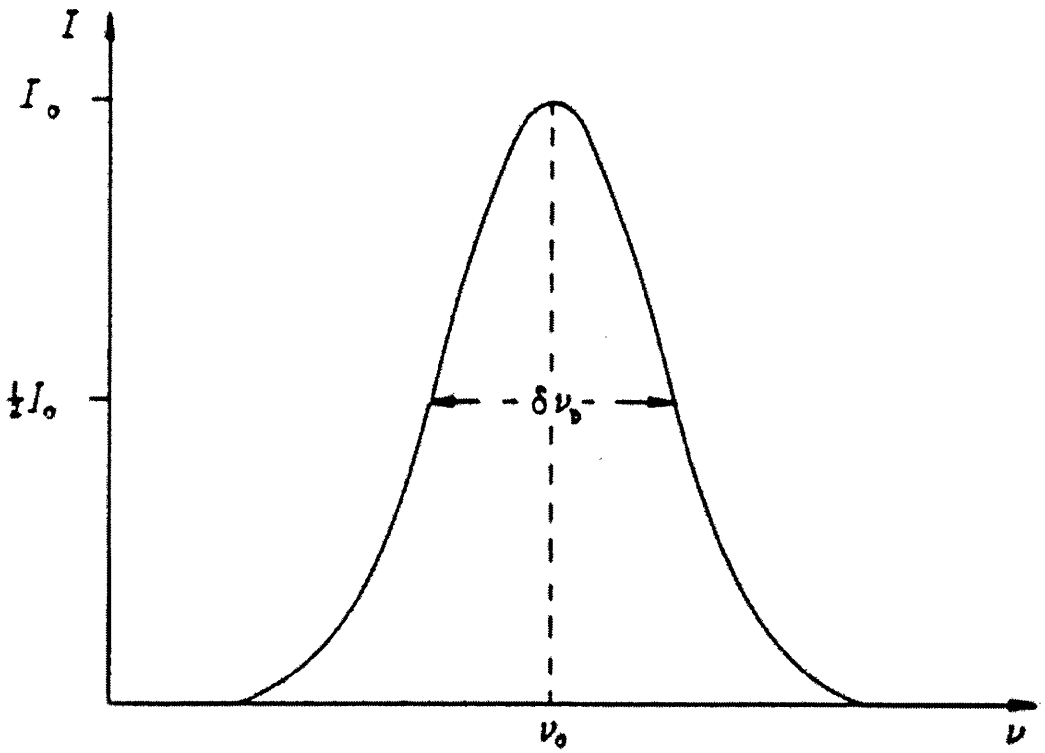


Figure 3.1 Doppler broadened, Gaussian line profile. (after Thorne, 1988)

Figure 3.1 shows the Doppler broadened, Gaussian line profile with the FWHM, $\delta\upsilon_D$ shown. In deriving the instrument function for an FPI instrument a Helium-Neon frequency stabilised laser may be sampled, providing profiles which are free from Doppler broadening. The profiles recorded during observations are compared to this instrument function, which has a correction applied to account for the difference in wavelength between the laser and observation wavelengths. This will alter the width of the instrument function profiles. The increase in width of the profile then provides information about the temperature at the altitude of the emitting species. This assumes that the emitting atoms have thermalized with their surroundings. In Section 1.7 this was shown to be true for the 630 nm emission of atomic oxygen from the thermosphere.

The emitting atoms may also have some component of a group velocity if there is a net wind acting on them. In this case the entire profile will be shifted in wavelength, to longer wavelengths if the wind causes net movement away from the observing station or to shorter wavelengths if the movement is towards the observing station. Measurements of the line of sight wind velocity can then be made by observing this shift in the recorded emission line profiles. In Section 3.2.3 the difficulty associated with determining the zero-shifted profile position, to allow calibration of wind velocities, will be discussed.

3.2.1 Precision of measurement

In making spectroscopic measurements with the FPI the finite width and shape of the emission line being studied provides a fundamental limitation to the accuracy of the location of the emission line (Hernandez, 1986). The flux of photons making up the measured signal will follow a discrete Poisson distribution (Parratt, 1961) and so the measured signal will also follow this distribution, scaled by the quantum efficiency of the detector. If the situation where an emission line is being sampled by a spectral scan at equidistant steps is considered, then the centre of the emission line is

$$\overline{\sigma} = \left\{ \sum_{i=-\infty}^{i=\infty} \sigma_i S_i \right\} \left\{ \sum_{i=-\infty}^{i=\infty} S_i \right\}^{-1} \tag{3.10}$$

where σ_i is the wavenumber at the ith step and S_i is the number of events measured at that step. The variance of determination of the centre of the line, $\delta\sigma_{\bar{\sigma}}$, is then given by

$$\delta\sigma_{\bar{\sigma}}^2 = \left\{\sum_{i=-\infty}^{i=\infty}(\sigma_i)^2 S_i\right\}\left\{\sum_{i=-\infty}^{i=\infty} S_i\right\}^{-2} \tag{3.11}$$

Gagne et al. (1974) have shown that for a Doppler broadened emission line of FWHM dg, the variance of determination of the centre of the line is given by the following expression

$$\delta\sigma_{\bar{\sigma}}^2 = dg^2[N\ln(2)]^{-1} \tag{3.12}$$

where N is the area under the profile, which is the total number of events counted by whatever detector is being used. In order then to increase the accuracy with which the centre of the profile is determined, and hence in the present application to determine the wind velocity information contained in the profile, this relation indicates that higher precision of measurement is more efficiently approached by reducing the line width than by increasing the number of photons counted.

Instrumental broadening will, in the light of the last equation, decrease the precision of measurement; however, the introduction of an extremely narrow bandwidth would produce a very low flux which would in turn negate the resolving power obtained by the extremely narrow bandwidth. The resulting compromise between instrumental broadening and luminous flux is termed the Luminosity Resolution Product (LRP) criterion (Chabbal, 1953).

Further consideration of the optimum operating parameters for the FPI is necessary if measurements of both winds and temperatures from the emission line sampled are to be made. Hernandez (1986) analysed the optimum instrumental parameter values for the FPI when making wind and temperature measurements and found that to produce the minimum uncertainty in one measurement would produce an increase in uncertainty in the other. The change in uncertainty for either wind or temperature however, when the other is optimised, was found not to exceed a factor of 1.1, so the instrument may be optimised for accuracy of either measurement without increasing the

uncertainty in the remaining measurement by more then 10%. While the LRP criterion can establish whether the choice of etalon aperture, etalon gap and plate reflectivity meet the required compromise between resolution and luminosity, it has no preference for the particular values used. When the measurement of wind and temperature in combination are taken into account however, these particular values require careful choice.

3.2.2 Determination of the zero Doppler shift position

Figure 3.2 illustrates diagrammatically the shifts in the position of the recorded spectral line as a result of making an observation on an emission layer with a constant velocity parallel to the Earth's surface. A constant uniform wind blowing from south to north would be observed as a red shift or blue shift depending on whether the direction of the observation was to the north or to the south of the station. The observations taken in this manner benefit from what is called Van Rhijn enhancement in that they sample a volume of atmosphere much greater than that sampled by looking directly overhead (see Figure 3.3). In obtaining the Doppler shift a zero-shift position for the peak needs first to be established. This however is not easy, as a calibration lamp would have to be used at the wavelength in question. Peterson et al. (1979) reported the development of such a lamp, however they are not in general use due to the difficulties encountered in production and operation. Two methods for determining the zero-shift position are

(a) to average the winds derived from a night of observations from opposite directions and averaging the wavelength position to produce a derived zero Doppler shift position. If there is a net wind over the observing station then a red-shift in one direction should be accompanied by a blue-shift in the opposite direction as demonstrated by Figure 3.2 . Averaging over the full range of observations should produce a good approximation to the zero wind position. The individual profiles can then be compared to this central wavelength position.

(b) to take observations looking directly overhead and assign to the centre of the profiles observed in this manner the zero Doppler shifted point i.e. the zero wind velocity point. This method, however, assumes that the vertical component of wind velocity is negligible and has been rejected after observations showed that there were often considerable fluctuations in the vertical component of neutral wind measured at

Kiruna and other high latitude sites (often of the order of 50 ms^{-1} under disturbed geomagnetic conditions, e.g. Aruliah et al. (1995)).

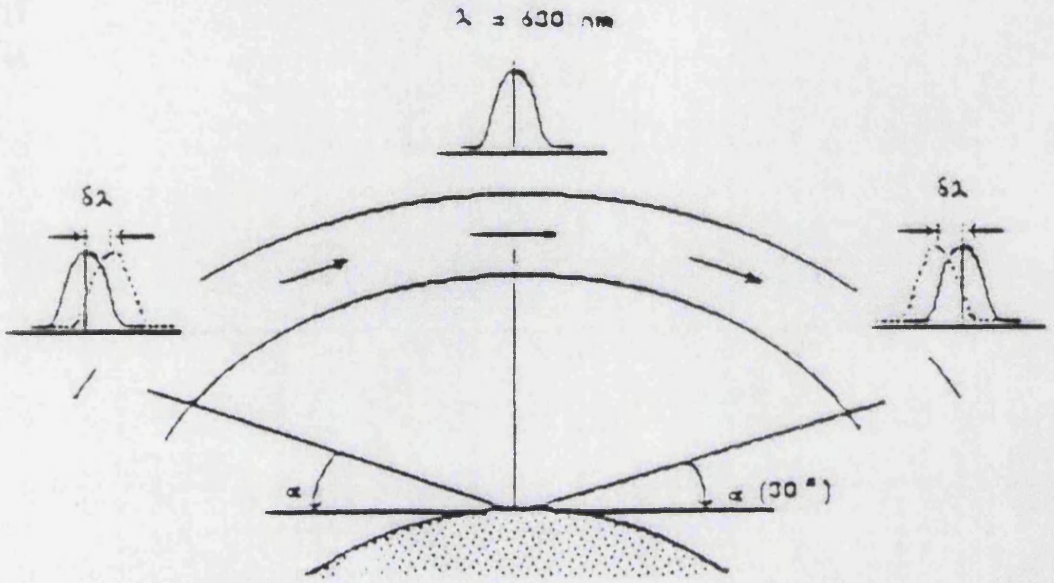


Figure 3.2 Apparent wavelength shift for different observation directions (after Doyle, 1987)

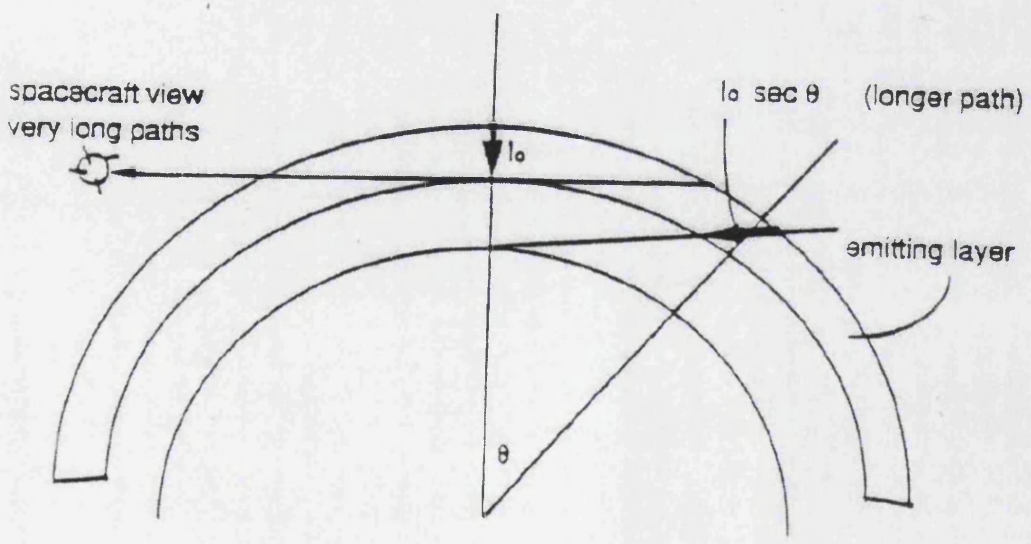


Figure 3.3 Van Rhijn Enhancement (after Killeen, 1990)

3.2.3 Scanning and stability with the FPI

In 1948 Jacquinot and Dufour introduced photoelectric recording to produce the Jacquinot and Dufour spectrometer which utilised a lens (L) to focus the etalon (F-P) transmission fringes to a plane containing an aperture (A), with a linear photoelectric detector behind the aperture, as shown in Figure 3.4.

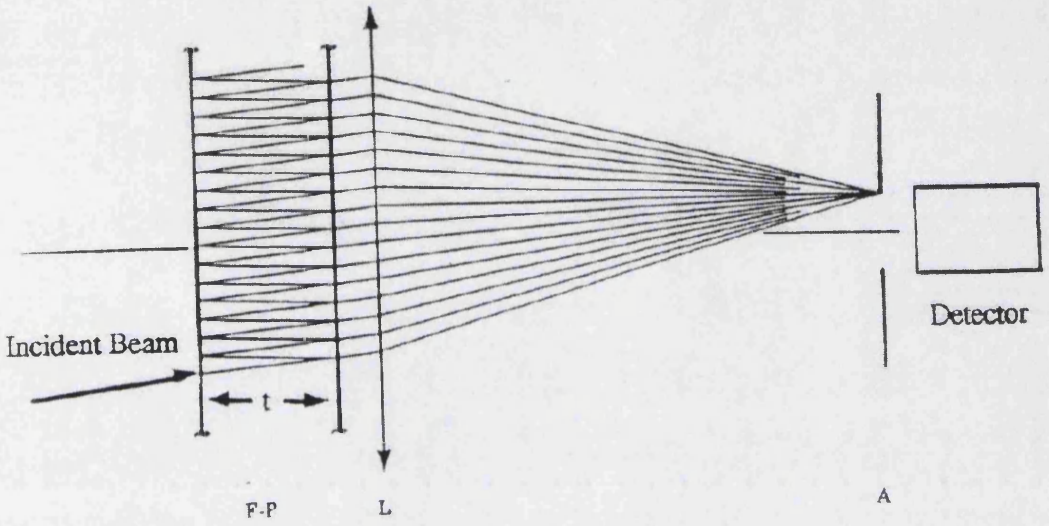


Figure 3.4 The Jacquinot and Dufour spectrometer

Operation of this type of spectrometer requires wavelength scanning. This means the aperture must be allowed to sample different portions of the wavelength spectrum. By writing the standard FPI equation

$$n = \frac{2\mu t \cos\theta}{\lambda} \tag{3.13}$$

it is shown that in order to achieve this change in the wavelength examined by the aperture (which will scan through the spectrum of the fringes and hence change the order n) there needs to be a change in one or more of the variables on the right hand side. Changing μ is known as index of refraction scanning, changing t is known as mechanical scanning while changing θ is known as spatial scanning.

Jacquinot and Dufour used index of refraction scanning when they produced the first scanning FPI (Jacquinot and Dufour, 1948), by changing the pressure of the air

between the plates of the etalon. This produced a consequent change in the refractive index of the air and hence achieved scanning of the fringe spectrum.

The limitations of this method are due to the non-linearity of the refractive index response to pressure changes, the variation in pressure needed to obtain an arbitrary refractive index change and the temperature sensitivity of the refractive index.

Hernandez (1986) has shown that the refractive index for air varies with pressure in the following manner

$$\left(\frac{\partial \mu}{\partial p}\right)_{T=15C} = (\mu_0 - 1)(1.3149\times10^{-3} + 1.626\times10^{-9}\, p) \tag{3.14}$$

where p is the pressure in Torr, μ_0 is the refractive index of standard air at 1 atmosphere and 15 degrees Celsius. It can be seen that the refractive index change relative to the pressure change is only non-linear to about 1 part per thousand million, and hence is negligible in most circumstances. Caplan (1975) derived the following formula to describe the effect of an increase in gas pressure on the gas refractive index in terms of the wavelength shift measured by the FPI:

$$\lambda - \lambda_0 = \left\{\frac{k\lambda_0}{T}\right\} P \tag{3.15}$$

where P is the pressure in bars, T is the temperature, k is a constant for a given gas, λ is the shifted wavelength and λ_0 is the wavelength at $P = 0$ and $T = 273.16$ K. This equation allows the pressure stability requirement to be calculated for obtaining winds to within a given accuracy.

Hernandez (1986) has also shown the temperature dependence of the refractive index at a given pressure can be shown to be

$$\left(\frac{\partial \mu}{\partial T}\right)_{p=760} = (\mu_0 - 1)\ 3.8753\times10^{-3}(1 + 0.003661(T))^{-2} \tag{3.16}$$

where T is in degrees Celsius. Caplan (1975) derived the following formula to describe the effect of an increase in gas temperature on the gas refractive index in terms of the wavelength shift measured by the FPI:

$$\frac{\Delta\lambda}{\Delta T} = \frac{-k\lambda_0 P}{T^2} \tag{3.17}$$

and using this equation found that for an FPI with an etalon gap of 7 mm and nitrogen gas used as the etalon medium, that the change in wavelength at 500 nm for a change in temperature, all other variables stable, is -0.00055 nm/C. This can be compared to the extent of a Free Spectral Range at this wavelength being 0.0178 nm. Thus for a change of 32°C in the temperature of the nitrogen, a full FSR would be shifted. Again, this allows the stability requirement, this time in terms of temperature, to be calculated to achieve a given wind speed accuracy.

Temperature changes also affect the spacers used between the plates of the etalon. In this case the change in wavelength caused by changing temperature can be described by the following equation:

$$\frac{\Delta\lambda}{\Delta T} = C_S\lambda_0 \tag{3.18}$$

where C_S is the coefficient of linear expansion for the spacer material. This equation may be used in order to set an upper limit on the acceptable value for C_S, and to choose an appropriate spacer material meeting this criterion to produce the desired wind accuracy.

3.3 FPI development

Hernandez and Killeen (1986) have described the history of optical measurements of the dynamical and thermodynamical properties of the upper atmosphere. These can be traced to 1923 when Babcock attempted to determine the Doppler width of the naturally-occurring O(^{1}D) radiation from the night sky with a Fabry-Perot spectrometer. An experiment carried out 11 years earlier in the laboratory by Buisson and Fabry had been designed to test the kinetic theory of gases by measuring Doppler widths as a function of temperature and Babcocks experiment was essentially an inverse of this. The results of the experiment allowed Babcock to infer an upper limit to the temperature at the height of the emission, now known to be near 96 km, in the lower thermosphere.

Following Babcock's measurement only limited quantitive information on the upper atmosphere was produced from other observations using instruments with photographic plates as detectors (e.g. Wark and Stone 1955; Cabannes and Dufay, 1956). The drawbacks inherent to these types of measurements included the low quantum efficiency and non-linear behaviour of photographic emulsions and also the Fabry-Perot silver mirror coatings employed were significantly inferior to the multi-layer dielectric coatings used today. These measurements however pointed to the potential for optical measurements of the upper atmosphere.

As demonstrated in Appendix A, Jacquinot and Dufour (1948) pointed to the radial symmetry of the Fabry-Perot spectrometer and later Jacquinot (1954) proved that for a given spectral resolution, it has the highest luminosity of those spectroscopic devices that depend on the interference of light for their operation. This led to the Fabry-Perot spectrometer being considered the instrument of choice for low-light-level, ground-based sensing of the upper atmosphere.

The ability of the Fabry-Perot spectrometer, used in conjunction with photoelectric detectors, to obtain quantitative information on upper atmospheric parameters soon led to many new measurements. An example was Armstrong (1958), who applied the method of Fabry and Buisson (1914) and Buisson et al. (1914), and observed the Doppler shifts of emitting species in the upper atmosphere to determine neutral winds. Hernandez and Killeen (1986) point out that, in part due to the simple and elegant physical principles upon which both the experimental technique and data analysis method are based, the use of Fabry-Perot spectrometers to determine temperatures and winds in the upper atmosphere became well-established following the work of Armstrong.

The APL at UCL has since the late 1970's operated a series of FPI's at high latitude sites in Scandinavia, usually observing the 630 nm emission of atomic oxygen, associated with an emission altitude around 240 km. These installations have typically been configured to allow remote, unattended operation throughout an observing season, which for these sites corresponds to the period from September to April when the sun is below the horizon for a significant length of time (> 4 hours).

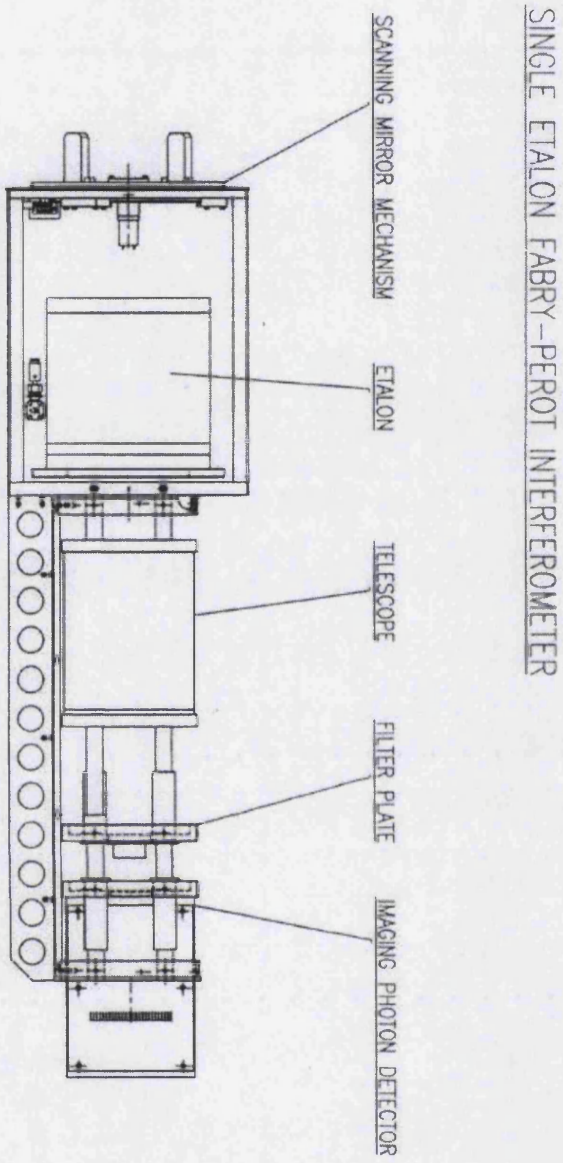


Figure 3.5 Early FPI schematic diagram.

A schematic diagram of a typical early FPI as used by the APL is shown in Figure 3.5. The instrument is hung vertically with the etalon at the top. A scanning mirror (not shown) is housed in a clear dome on top of the instrument building. This mirror allows the viewing direction to be changed. A calibration lamp is used to take calibration images during observing periods, allowing any instrumental faults to be detected. A fixed mirror is used in combination with the scanning mirror to take zenith observations. The optical bench from this design contains a Cassegrain telescope assembly to focus the image onto

the detector. The bench also contains a filter which is centred on the wavelength of interest. The detector is attached to the end of the bench.

3.3.1 Fabry-Perot Etalons

The etalons are constructed from plates of fused silica (Spectrosil B) and mounted within sealed temperature controlled enclosures, while the plate parallelism and separation are maintained by sets of three Invar spacers of nearly identical length. Temperature control (to better than 0.1°C) is essential to prevent the optical path within the etalon cavity changing either due to the variation of the refractive index of the medium within the cavity or due to the temperature dependence of the various mechanical elements. Possible temperature dependent processes include (Killeen et al., 1982)

(1) thermal expansion of the spacer material defining the gap
(2) change in refractive index of the medium comprising the gap
(3) thermal expansion of reflective coatings
(4) change in the reflective indices of the reflective coatings
(5) change in the internal stress of the multi-layer coatings leading to a change in the surface quality of the end plates
(6) outgassing of absorbed material from coated surfaces
(7) distortion of spacer/end plate bonds due to the mismatch between the coefficients of expansion.
(8) Change in the force on the etalon due to its mounting arrangement (axial or radial).
(9) distortion of interference ring pattern due to "ghost" fringes from internal reflections between surfaces.

Owing to the optical contacting used throughout the device, the etalon is stable against mechanical shock and vibration for most laboratory and field applications. Invar is chosen as the spacer material due to its very low coefficient of thermal expansion. Figure 3.6 shows a fixed etalon, this was the actual etalon used in the FPI flown on the DE-2 satellite.

The APL at UCL has also been involved in using Fabry-Perot etalons for space-flight applications where the spacers have been bonded using cyanoacrylate adhesives. This bonding technique enables the etalon to survive extreme stresses without

measurable degradation of the optical performance under normal operating conditions (Rees et al., 1981).

The coatings used in modern Fabry-Perot etalons are multi-layer dielectrics which provide the required spectral characteristics for each application. As seen in Section 3.2.1 the choice of reflectivity and etalon gap need to be optimised for specific applications. The most recent APL etalons are coated to provide reflectivities of 87.5 % +/- 4% from 600 nm - 900 nm and are specified to a flatness of $\lambda/150$ or better over this range. The etalon gap is chosen to be 10 mm or 14 mm, ensuring that the detectors in current use can produce images from which accurate wind speeds may be determined.

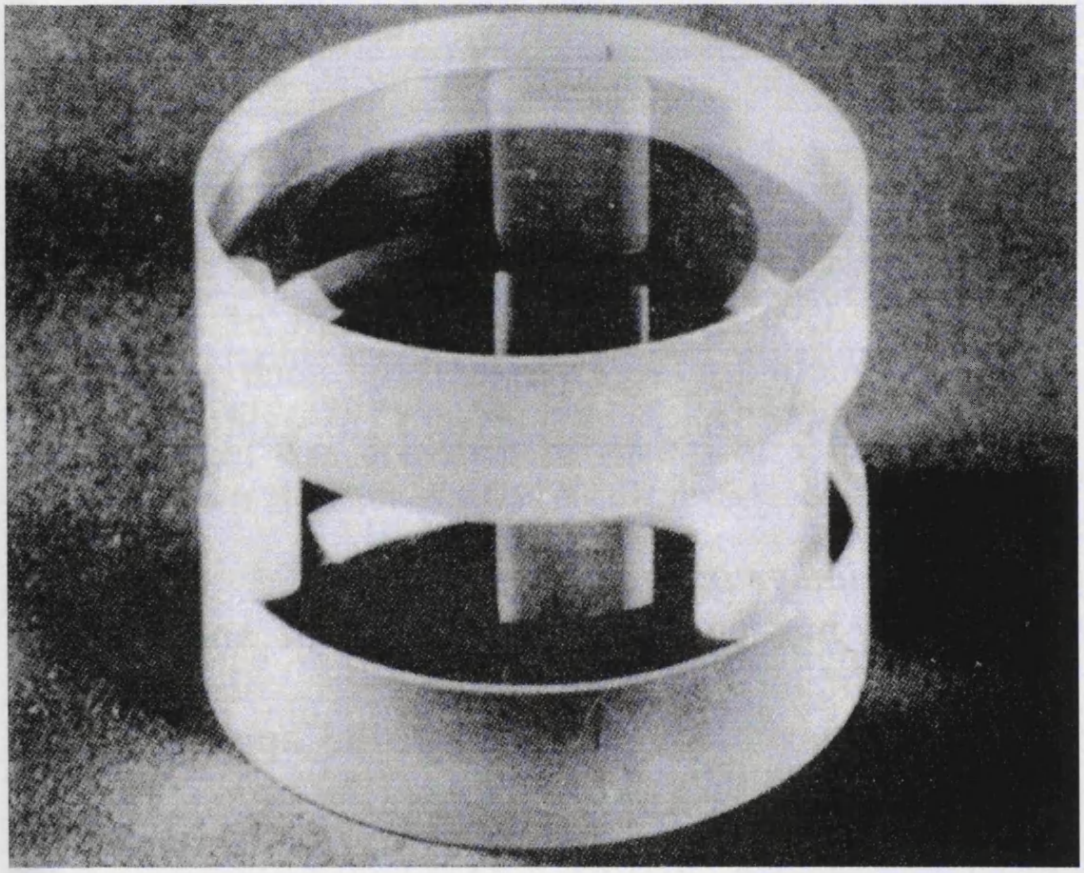


Figure 3.6 Fabry-Perot etalon (after Killeen et al., 1982).

Capacitance stabilised etalons (CSE), which have etalon gaps that may be controlled via stacks of piezoelectric transducers, stabilised by a capacitance micrometer servo-loop, have also been developed for specific applications (e.g. Rees et al., 1981). This enables mechanical scanning to be performed. The capacitive elements are integral

features of the three spacers, which also incorporate the piezoelectric transducers. A schematic diagram of a CSE is shown in Figure 3.7 .

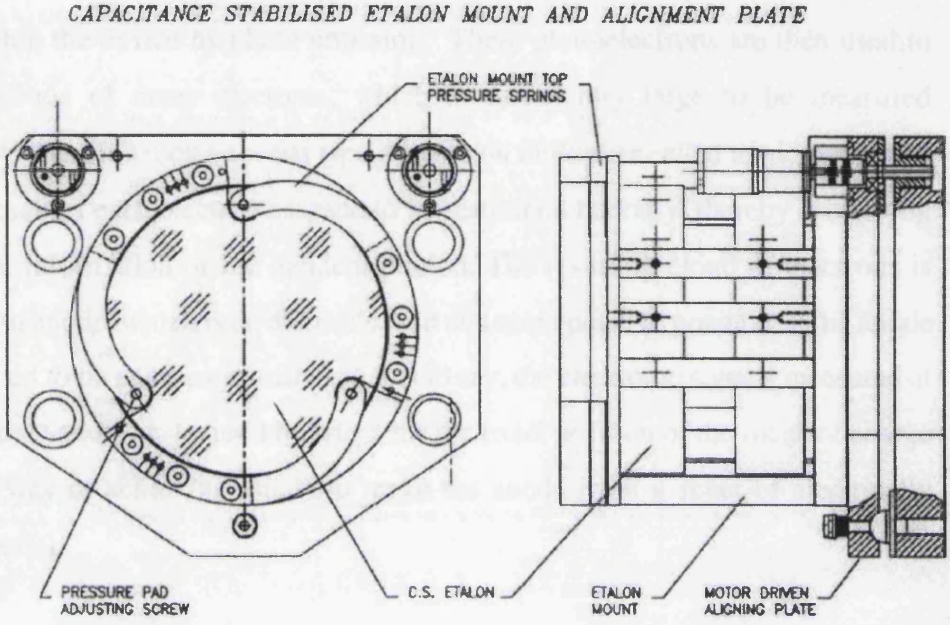


Figure 3.7 Schematic design of a capacitance stabilised etalon in its mount.

3.3.2 Detector development

Until recent years it has been difficult to take advantage of the full light-gathering potential of the Fabry-Perot etalon. The extremely low light levels involved require the use of a detector which is sensitive to individual photons. As mentioned in Section 3.3 the use of photographic plates in early studies was superceded by the use of conventional photomultiplier tubes as the detector (e.g. Armstrong, 1956) to measure atmospheric Doppler shifts. Such devices, similar to the Jacquinot and Dufour spectrometer referred to in Section 3.2.3, are not capable of imaging, and can only measure the intensity of one part of the fringe pattern at a time. They are usually placed on the optical axis of the instrument and a sequence of measurements is made whilst the fringe pattern is scanned through the collecting area. Because such detectors have no spatial resolution, any increase in the size of the collecting area results in a blurring of the measured profile, and the aperture would therefore be commonly restricted to a pinhole.

3.3.2.1 IPDs

In the 1970's the APL embarked upon a programme of development of a new type of detector, known as the Imaging Photon Detector (IPD), shown in a schematic diagram as Figure 3.8. This device is similar in concept to a photomultiplier tube in that it is sensitive to individual photons of light. The incident photons generate individual electrons within the device by photo-emission. These photoelectrons are then used to initiate a cascade of many electrons, which is sufficiently large to be measured electronically. The IPD uses a special type of electron multiplier called a micro-channel plate. This enables each electron cascade to be restrained laterally, thereby preserving the positional information of the incident photon. The resulting cloud of electrons is collected by an anode, which is an electrode held at a more positive potential. The anode can be designed to be position sensitive, that is to say, the electronic signals measured at the edges of the anode can be used to determine the exact position of the incident charge cloud. One way of achieving this is to make the anode from a sheet of electrically resistive material.

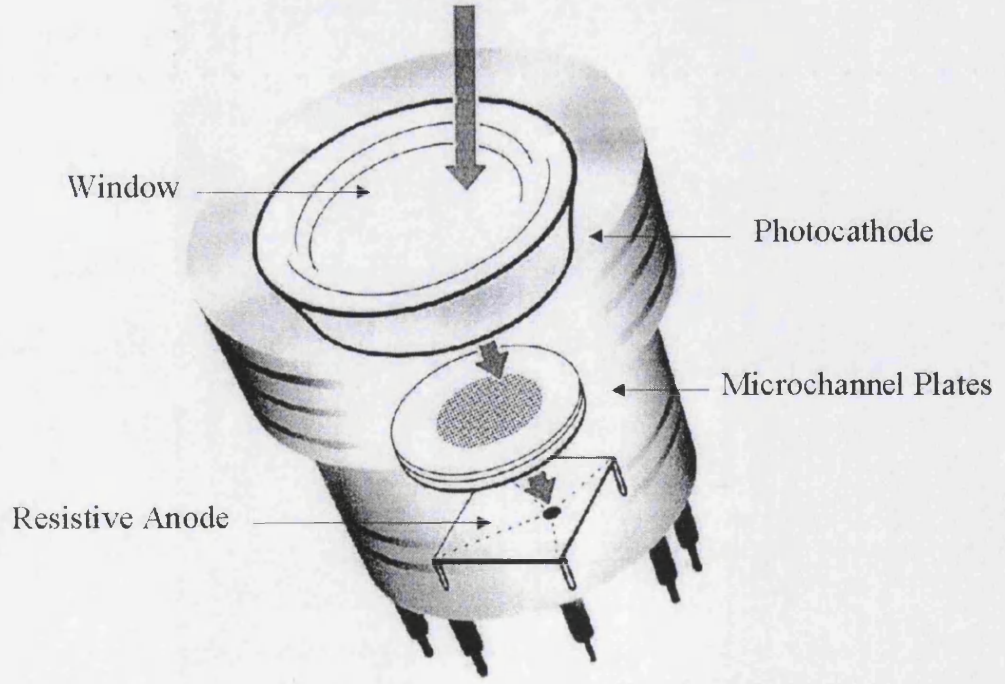


Figure 3.8 An IPD detector schematic showing the path from photon incidence on the photocathode, through the micro channel plates to detection at the anode.

IPD development allowed full Fabry-Perot rings to be recorded with high time resolution in FPI instruments (McWhirter et al., 1982). These detectors allowed non-scanning etalons to be used to determine line widths and shifts. Their advantage was the fast readout times for the image but their main disadvantage was the non-linearities in the image quality due to defects in the resistive anodes.

These IPD devices have been used since the early 1980's on various instruments at various locations across the world. The devices lose quantum efficiency gradually with age and are at risk of destruction through over exposure to high light levels, however several such devices have performed satisfactorily for over a decade in individual instruments. At present the IPD is being used as the detector on the APL FPI at Kiruna observing the oxygen red line, with another instrument operated by the Space Physics Department at University of Latrobe, Melbourne, Australia.

3.3.2.2 Charge Coupled Devices (CCD)

A CCD is an ultra-sensitive light detector made primarily of silicon and manufactured using similar technologies to those used in making silicon chips. One of the basic properties of silicon is its very high sensitivity to light; it can easily absorb photons of light that are incident on its surface. When a photon of light is absorbed, a single electron is released that is free to move about the silicon crystal lattice structure. However, CCDs are specially designed to store these generated electrons and prevent them from wandering around the lattice. In this way a pattern of electrons is built up in the CCD that directly corresponds to the pattern of light received. The resultant electron charge pattern can then be read out from the CCD using analogue electronics and finally digitised to give a very accurate digital representation of the light image that has fallen on the CCD.

Until the 1980's however, commercially available CCD chips suffered from inherent noise levels that prevented their use as a "bare" device for observing airglow emissions because of the extremely low light levels involved and the inability to discriminate the signal from this noise. A solution was devised which used a CCD as the final detection stage but coupled a cathode to the front end to provide sufficient signal enhancement to allow the CCD image to be taken out of the noise-limited regime. The APL has used such a hybrid device, called an intensified CCD (ICCD), at Kiruna for a number of years with success, however it has a significant drawback in terms of the

persistence of the phosphor used on the cathode. This produces an "afterglow" effect whereby some remnant of a previous image will be recorded on any image taken directly afterwards. The effect led to the need for a gap between successive integrations in normal operation, negating the time resolution improvements introduced by the ICCD technique.

More recently bare CCD technology has been used in FPI instruments (Niciejewski et al., 1994). The results showed that the sensitivity of these devices was now sufficient to allow their deployment in ground-based remote sensing of thermospheric airglow applications. The bare CCD devices used by the APL in recent instruments have been manufactured by Astrocam Ltd. (utilising CCD chips supplied by EEV Ltd.). The specification of this CCD chip is shown in Table 3.1.

Table 3.1 Super-MPP CCD specification

Pixel Array Format (Horizontal x Vertical)	770 x 1152
Sensitive Area (mm)	17.3 x 25.9
Pixel Size (microns)	22.5 x 22.5
Typical Full Well Capacity (electrons)	210,000
Typical Read Noise (e^-) at 40 kHz	4
Typical Dark Signal (e^-/pixel/sec)	0.01 - 0.1

The most important figures of merit for the chip in the application of observing airglow with an FPI relates to the dark noise level (dark current), generated within each pixel by thermionic emission. The chip uses Multi-Phase Pinned (MPP) technology to achieve these low noise levels which reduces dark current during long exposure times but reduces the capacity of the pixels of the CCD to store charge, known as the full well capacity. In its super-MPP guise this technology reduces dark current by a factor of 1000 compared to conventional non-MPP devices. Due to the low light levels observed in the airglow observing FPI application, the full well capacity is not as important a requirement

of the CCD as the low dark current. To fulfil this requirement the device also employs liquid cooled peltier stacks to cool the chip to around -50°C.

3.3.3 Future FPI development

A number of different strategies have been postulated for improving the scope of FPI measurements in recent years. Some novel developments under consideration are the use of multiple etalons to extend measurements into daytime and the use of all-sky imaging to produce accurate wind vectors across the entire field of view.

The use of multiple etalons has been tested in the past for daytime measurements, their role being to discriminate the emission of interest from the overwhelming solar continuum background. However difficulties encountered include the complexity of optical alignment and the linearity of the detector response under dramatically dynamic illumination changes. For all-sky imaging the detector linearity considerations are also limiting but the time resolution required to record accurate profiles from each resolved sector of the image need to be considered.

A practical example of the application of the all-sky philosophy has been demonstrated recently by Conde and Smith (1997) and it is expected that this development will continue. A major advance however in this regard has been the recent development of cheap bare CCD detectors with low light level application capabilities, specifically the super-MPP type chips.

3.3.4 Current APL facilities

The APL currently has FPI's in operation measuring thermospheric neutral winds from three different sites: Kiruna, Sweden; Longyearbyen on Svalbard, Norway; and RAL in Oxfordshire, UK. Each of these sites has unique properties in terms of the physics at play within the thermosphere observable from each site.

The RAL site is at mid-latitude where the winds are expected to follow a strong diurnal pattern with poleward winds during the day, weak winds near dawn and dusk, and strong equatorward winds peaking near local midnight. As such it represents an opportunity to investigate basic thermospheric conditions without the complication of regular perturbation from the anomalous behaviour of the electrojet observed at low latitudes, or the influence of geomagnetic activity as seen in the higher latitudes.

At Svalbard the site is usually within the polar cap as defined by the auroral oval. One of the most important physical processes driving the thermospheric response to a particular level of geomagnetic activity is the polar ion convection driven by the magnetospheric field. The polar ionosphere is connected to the magnetosphere via the geomagnetic field lines and hence the ion convection is essentially driven by the interaction of the geomagnetic field lines with the solar wind. In general the neutral wind in the polar thermosphere is driven by two forces of quite different origin: the general circulation of the atmosphere due to the solar heating, and the general circulation of the magnetosphere due to the solar wind (Hargreaves, 1979). The Svalbard FPI is therefore well positioned to investigate the dynamic interplay of the competing forces in the thermosphere at this latitude. Recently both the EISCAT Svalbard Radar (ESR) and a digital ionosonde have been installed on Svalbard close to Longyearbyen and these will enable intercomparisons to be extended to this site in the future.

At Kiruna geomagnetic activity dictates whether the site is inside or outside the auroral oval. This leads to a wide range of thermospheric flow patterns being observed reflecting the diverse and dynamic nature of the forces in action. In combination with the Svalbard FPI it can give a clear picture of the high latitude thermospheric wind patterns and their response to geomagnetic activity and energetic particle precipitation from within and outside the auroral oval.

3.4 Mid-latitude site: Rutherford Appleton Laboratory

From Chapter 2 we have seen that the emphasis for research within the thermosphere has in recent years been on the complex dynamics at play in high latitude and low latitude regions where the interaction of the ionosphere and thermosphere is subject to major external influences. Much work can still be carried out at mid-latitudes with relevance to the current efforts in the field however, and at the same time expand the database of empirical measurements. This needs to be continued to allow changes in the climatology of the thermosphere to be investigated and the long term situation to be derived. The work of Buonsanto (1990, 1991), Hedin et al. (1994), Titheridge (1995b), Miller et al. (1997) and Buonsanto and Witasse (1999) have developed the understanding of this climatological behaviour at mid-latitudes while pointing to the need for more measurements, especially where co-ordination of different techniques is possible, and for

better understanding of the forces driving this behaviour through more detailed theoretical modelling work.

The RAL site has special significance in this respect as it continues the longest running database of ionosonde measurements in the world. In 1993 the Slough ionosonde was replaced by the RAL digisonde after a lengthy process of cross-comparison. The Slough database extends back to 1932 and has been used for numerous studies of the long-term behaviour of the ionosphere (e.g. Miller et al., 1997; Field and Rishbeth, 1997).

3.4.1 RAL FPI

The RAL FPI is a new facility which is still in the process of full installation but will enable the APL to test upgrades and developments for all of its FPIs at a local site with easy access. It is the newest in the series of instruments used by the APL to measure thermospheric winds and is installed in a purpose built, air conditioned and temperature controlled hut on top of a three storey building on the RAL site. First data was taken in February 1998 and the instrument was put into regular operation in April 1998. Unfortunately the number of nights with significant data is small as there have been a number of extended periods of in-operation with the instrument since 1998, some due to mechanical failures but also due to the need to support the remote high latitude sites using items from the RAL FPI. In combination with the requirement for clear skies this has restricted the database available.

Figure 3.9 shows a schematic diagram of the instrument. The important features of this implementation of the Fabry-Perot design are the use of the bare CCD detector and one of the newer generation of lens based optical benches. This bench replaced an older Cassegrain telescope optical bench in the winter of 1998/99 to enable enhanced illumination of the CCD, leading to higher time resolution. It has also been necessary to provide baffling around the bottom of the dome to prevent local light sources such as car park lights from polluting the FPI data.

The hut also contains the ancillary equipment needed to maintain the operation of the instrument. This includes a water cooler and pump to provide the necessary coolant flow to the peltier stack housed in the CCD detector head, the control electronics used to position the mirror and trigger the calibration lamp and also a PC to coordinate the control aspects of the instrument and run the data acquisition software. This software has been developed at APL over many years and is updated to allow integration of the latest

detector technology into the existing instrumentation. The PC also allows remote control and data transfer from the instrument via the internet.

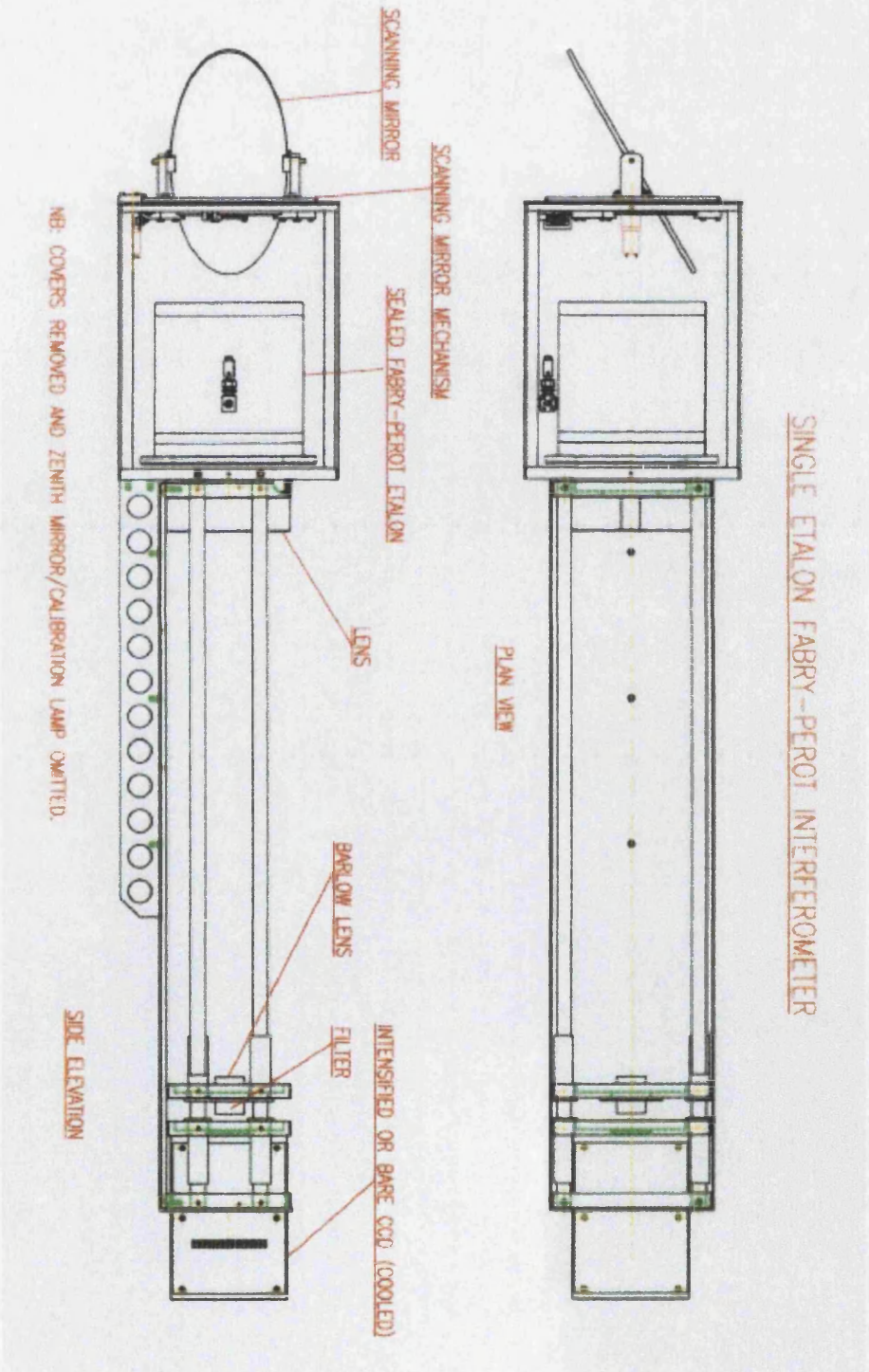


Figure 3.9 Schematic diagram of the RAL FPI

3.4.2 RAL Digisonde

In 1993 RAL purchased a Digisonde from the University of Lowell in order to replace the analogue ionosonde system which had been in operation from the 1930's at Slough in England. A Digisonde (DGS) is a pulse-amplitude ionospheric sounder, designed both as a monitoring and research tool. The instrument has the capability of measuring many ionospheric parameters such as time-of-flight group range, amplitude, Doppler velocity, angle-of-arrival, wave polarisation and wave-front curvature. At RAL the system employs a spaced-array of four receiving antennae. A real-time Fourier transform of the received signal gives the amplitude and Doppler estimate for each source echo, and interferometry is used to determine the source position.

The DGS is designed to operate in one of two modes: ionogram and drift. In ionogram mode, the reflection range and echo amplitude (accurate to 0.5 dB) are measured over a selected frequency range. Drift mode, in which a single frequency is chosen, is used for time-series recordings and allows measurements of the drift velocity to be made from these higher time-resolution data.

The frequency range for the DGS is 0.4-40 Mhz. The echoes are received by crossed-loop antennae that enable the O and X-mode echoes to be distinguished. The technical criticisms of the instrument centre on the pulse compression techniques employed and the large bandwidth. The lack of adequate filters in the transmitter stage can give rise to considerable interference to other, nearby instruments. At RAL the transmitter is situated apart from the main buildings and shielded by a small hill.

An important advantage of the DGS is the facility for automated control. This means that DGS systems can be linked in a chain covering a wide geographical spread and programmed through a central computer linked into either a standard telephone network, or over the Internet. Digisondes have been produced in large numbers and operate as part of a world-wide network. RAL operates another two ionosondes in this manner at remote locations at Lerwick in the Shetland Islands and Port Stanley on the Falkland Islands. Automatically scaled parameters such as *foF2*, *foE* and *M3000F2* are available from these instruments on the Internet within days, while the more accurate IIWG format data is made available on the World Data Center – C1 (WDC-C1) web pages at RAL within a couple of months of observations being taken.

3.4.3 Combined Measurements

At mid-latitudes both FPI and DGS instrumentation can suffer from low signal to noise conditions under normal circumstances due to low airglow emission intensities for the FPI and low electron densities for the DGS. While this does not preclude measurements being taken it decreases the number of times when combined data is available. Coupled to this are the problems specific to an FPI - DGS comparison whereby the FPI operates during the hours of darkness, while the DGS works least well around local midnight due to the low ionospheric electron densities.

In order to compare the datasets of the instruments at RAL it is important to understand the relevance of the measurements. The FPI can only derive vectors of thermospheric winds when it is operated on clear nights. Its viewing geometry means that the derived meridional winds are associated with a thermospheric horizontal area of 240 km radius. In comparing these winds to those derived from the RAL digisonde it should be noted that an assumption is made that there is no divergence in the wind field over this area. The two instruments are located approximately 200 metres apart at ground level and therefore their ground displacement is negligible compared to the assumptions used in their comparison.

Hence while the FPI and DGS at RAL have been taking coincident measurements since February 1998 very few nights have produced good datasets from each instrument for comparison. Titheridge, (1995b), has pointed to the importance of the use of ground-based radio measurements of *hmF2* to obtain neutral winds, because of the confinement of most optical methods to nighttime and their vulnerability to adverse weather conditions. The method has the potential to vastly increase the wind data base because ionosondes have been making routine measurements on a global scale for many years. However, because of the lack of coincident radio and optical measurements, few validation studies have been carried out.

One example is the work of Dyson et al. (1997) in the southern hemisphere, where it was found that, for the mid-latitude site at Beveridge, near Melbourne, the HWM model failed to match the amplitude of the midnight winds from FPI measurements. These amplitudes were well matched by the corresponding winds derived from Digisonde measurements of *hmF2* at the site however. An example of the comparisons possible so far at RAL is shown in Figure 3.8, in which it is seen that the Digisonde and FPI winds agree well while the HWM output again fails to match the midnight amplitudes. Also

shown is the output generated using the IRI model as a source of *hmF2* values for the MWM. Again this fails to match the wind amplitudes from the measured parameters.

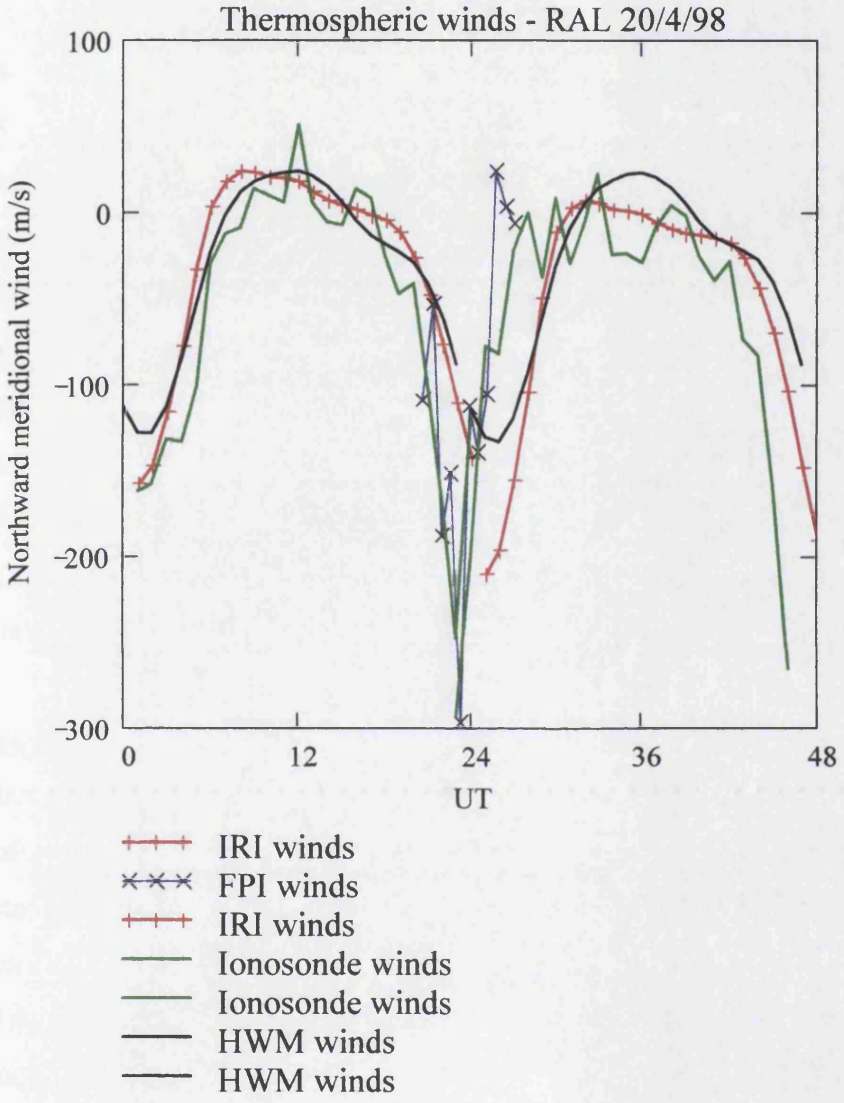


Figure 3.10 Comparison between RAL Digisonde (denoted Ionosonde winds) and FPI, with IRI and HWM model predictions for 20/4/98

It is envisaged that the database of coincident measurements at RAL will become a valuable tool for mid-latitude thermosphere-ionosphere studies.

3.5. High latitude site: Kiruna/Tromso

The diurnal precession of the geomagnetic pole around the geographic pole causes the ionospheric convection cells to oscillate in geographic latitude. This oscillation of the convection cells moves Kiruna in and out of the auroral oval. During the midday sector Kiruna is equatorward of the auroral oval. As well as the diurnal oscillation the exact size of the oval is always varying, depending on the degree of geomagnetic activity. During very quiet periods, the oval contracts poleward of Kiruna. During very disturbed periods the oval expands equatorward of Kiruna.

As a result of the complexity of the ionospheric and thermospheric behaviour , this area of northern Sweden and Norway is well furnished with instruments making measurements of thermospheric and ionospheric parameters, including radars, optical instruments, magnetometers etc. In particular the Kiruna FPI and EISCAT UHF radar, located at Kiruna (67.4°N, 20.4°E) and near Tromso (69.6°N, 19.2°E) respectively, have been used for combined measurements through the EISCAT INDI experiment (Davis et al., 1995) in the past. The present study combines these instruments along with other neutral wind sources to extend the scope of the comparisons possible in this region.

3.5.1 Kiruna FPI

The APL at UCL has maintained an FPI at Kiruna from the early 1980s. The majority of the data have been recorded using a Cassegrain bench assembly with 150 mm aperture etalon and IPD detectors, although more recent data have been recorded using an ICCD detector and lens-based optical benches. These instruments were similar to those presented as examples earlier in this chapter. Early data was taken from a site beside the Esrange rocket range, 50 km from Kiruna, before purpose built optical observatories were built on the roof of the Insitute for Space Physics (IRF) at Kiruna, which provided the APL with two observing rooms for its FPI's from 1982 to late 1999.

A particular advantage of having these observing rooms at the IRF has been the ability to test developments side-by-side with existing instrumentation. One such comparison was undertaken between the IPD based FPI and ICCD based FPI. This is demonstrated in Figure 3.11 which shows the northward line-of-sight thermospheric winds determined from either instrument. The superior time resolution of the ICCD device enables a much clearer delineation of a feature with large wind gradients between 2200 and 2300 UT observed in both datasets.

At the end of 1999 the two FPIs were rehoused at a new site near Esrange in temporary observing huts, which are modified versions of those used by the Auroral Imaging System (ALIS) experiment. This site, named KEOPS for Kiruna-Esrange Optical Site, will be the site of a new purpose built optical observatory.

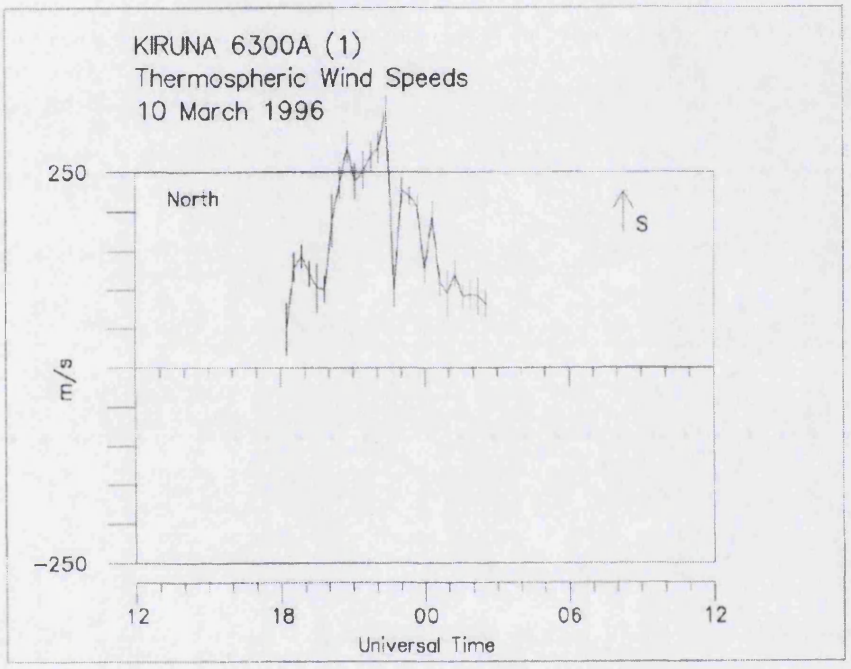


Figure 3.11 (a) Derived thermospheric winds from IPD FPI.

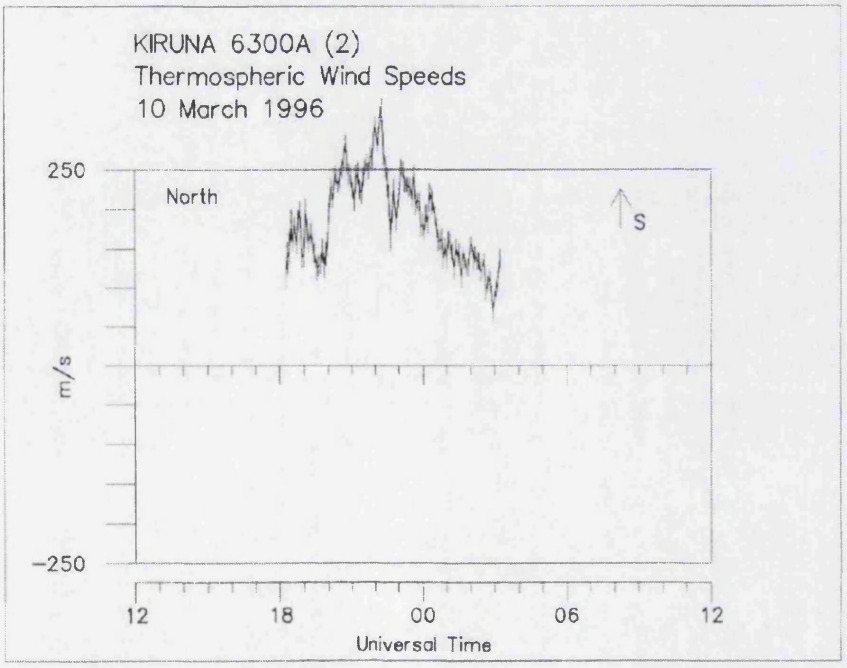


Figure 3.11 (b) Derived thermospheric winds from Intensified CCD FPI

3.5.2 EISCAT System

The EISCAT incoherent scatter radar is managed by the EISCAT Scientific Association, an organisation which was established in 1975 by the scientific research councils of the United Kingdom, Germany, Finland, France, Norway, and Sweden – known as the EISCAT Associates. Their aim was to conduct research on the upper atmosphere, ionosphere and aurora using the incoherent scatter technique. In 1996 Japan joined this group as a full member. The system first began operations in 1981.

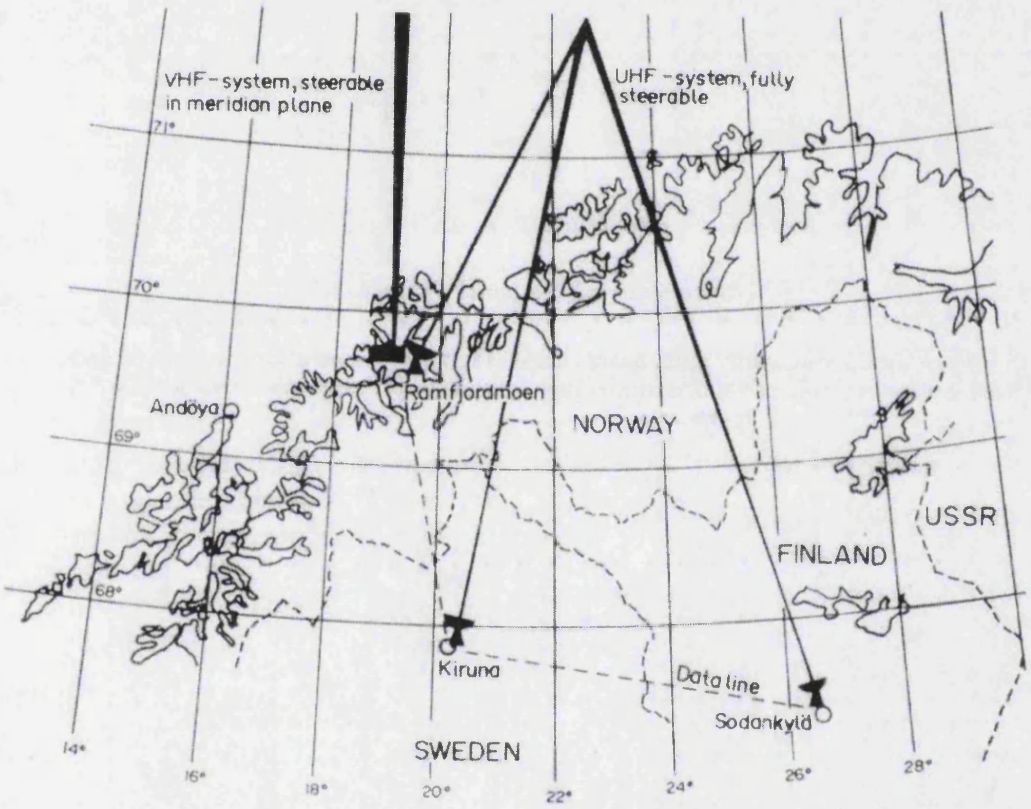


Figure 3.12 The Eiscat System sites in Northern Scandinavia.

Figure 3.12 shows the locations of the experimental sites for the EISCAT system in Northern Scandinavia. The transmitting site is at Tromso, which is also a receiver, and the receiving sites are at Sodankyla and Kiruna. The system consists of two independent radars, a tristatic UHF radar operating at 931 MHz and a Monostatic VHF radar operating at 224 MHz. Rishbeth and Williams (1985) describe the system in detail as do the series of EISCAT technical reports.

The UHF and VHF radars share a single high-voltage capacitor bank which provides close to 90 kV for the klystrons during normal operation. For the experiments contributing to this study the UHF radar had a single VARIAN klystron which operated at a central frequency of 931.5 MHz, with a bandwidth of 8 MHz, a nominal peak power of 2 MW, and a maximum duty cycle of 12.5%, giving an average power output of 250kW. There is a minimum of at least 1 ms between periods of transmission and pulse duration ranges from 1 μ s to 10 ms. Radio-frequency power is actually transmitted for less time than the quoted duty cycle as a certain length of time is required after switching on the beam current before a steady enough signal is established. The UHF klystron is being replaced with a pair of new klystrons and is due to resume operations in 2000.

Each of the three antennae in the UHF system is a fully steerable 32 m Cassegrain-feed, parabolic dish antenna. These are each steered by motorized azimuth and elevation drives, with an azimuth range of 270° either side of geographic north. They have an elevation range of between 2° and 100° and have a maximum scan rate of 80° min^{-1}.

For the antennae at the remote sites there is a need for a very high degree of precision in antenna pointing. Lehtinen and Turunen (1981) demonstrated that the r.m.s. error in pointing accuracy can be reduced to less than 0.02° by applying a correction to each pointing instruction. Together with a 1 μ s timing precision this gives the accuracy needed for tristatic-velocity measurements.

EISCAT experimenters interact with their experiments through the EISCAT Real-Time Operating System (EROS), within which a control language called TARLAN (Transmitter and Receiver Language) is used to program the radar controllers. This in turn controls the receiver-protect circuits at Tromso, the high-voltage to the klystrons, the RF signal, receiver switches, calibration signal and instructions to the correlator. Each EISCAT site is also equipped with a caesium-standard clock to provide frequencies of 1 MHz, 5 MHz and 10 MHz to an accuracy of one part in 10^{11}. The radar controller accepts 1 μ s timing signals from the clock output to direct the transmitter and receivers.

EISCAT experiments implement control through a high-level language called ELAN (EISCAT Language) run on Norsk Data computers. Programs written in ELAN control the antennae, polarisers, attenuators, filters and ADCs, etc. They also provide control of the correlator and radar controller, setting the sampling rate, data-integration times and organising storage of the data. This regime is under review however as the

operation of the ESR and its more recently developed suite of control software has prompted suggestions for a similar system to be implemented at EISCAT, which is currently being undertaken.

EISCAT operations are divided equally between Common Programmes (CP) and Special Programs (SP). Common Programmes are operated routinely while Special Programmes are run by the EISCAT Associates for particular research applications. There are four UHF and three VHF Common Programmes. These are run for a total of 750 hours every year. There are also a number of Unusual Programmes (UP) which can be run at short notice during "unusual" geophysical conditions such as extensive energetic-particle precipitation, and for auroral studies.

The four UHF CPs are:

CP-1 : with a fixed, field-aligned Tromso beam, the remote sites scan between the E and F region. This experiment provides the high time resolution needed to study rapid changes in auroral phenomena and altitude coverage is from 85 – 600 km.

CP-2 : a 4-position scan covering field-aligned, vertical, North and East, with a 6 minute cycle time. Altitude coverage is 85 – 600 km. An advantage over CP-1 is that the plasma velocity is measured at four different positions, however this can cause problems if the ionosphere changes significantly over time scales less than the scan time.

CP-3 : a 17-position meridional scan with 30 min cycle. This experiment covers a wide latitude band from 62 – 76°N. Altitude coverage is from 140 – 950 km and the remote sites follow the scan in the F region.

CP-5 : a rapid latitudinal scan combined with a field-aligned beam which is essentially a cut down version of CP-3 providing a faster cycle time.

The three VHF CPs, designed for D-region studies and high-altitude studies above 1000 km are:

CP-4 : a high latitude programme (up to 75°N) with a split-beam radar. This experiment is designed for the study of plasma flows to the north of the auroral zone.

CP-6 : a D-region programme covering a restricted range from 76 – 133 km at a resolution of 1 km using pulse-to-pulse correlation to achieve the very-high spectral resolution nesessary at these heights.

CP-7 : a Topside programme. This experiment uses long-pulse codes to observe the topside ionosphere at the highest altitudes, up to 2500 km, and is used to study the plasma flow between the ionosphere and magnetosphere.

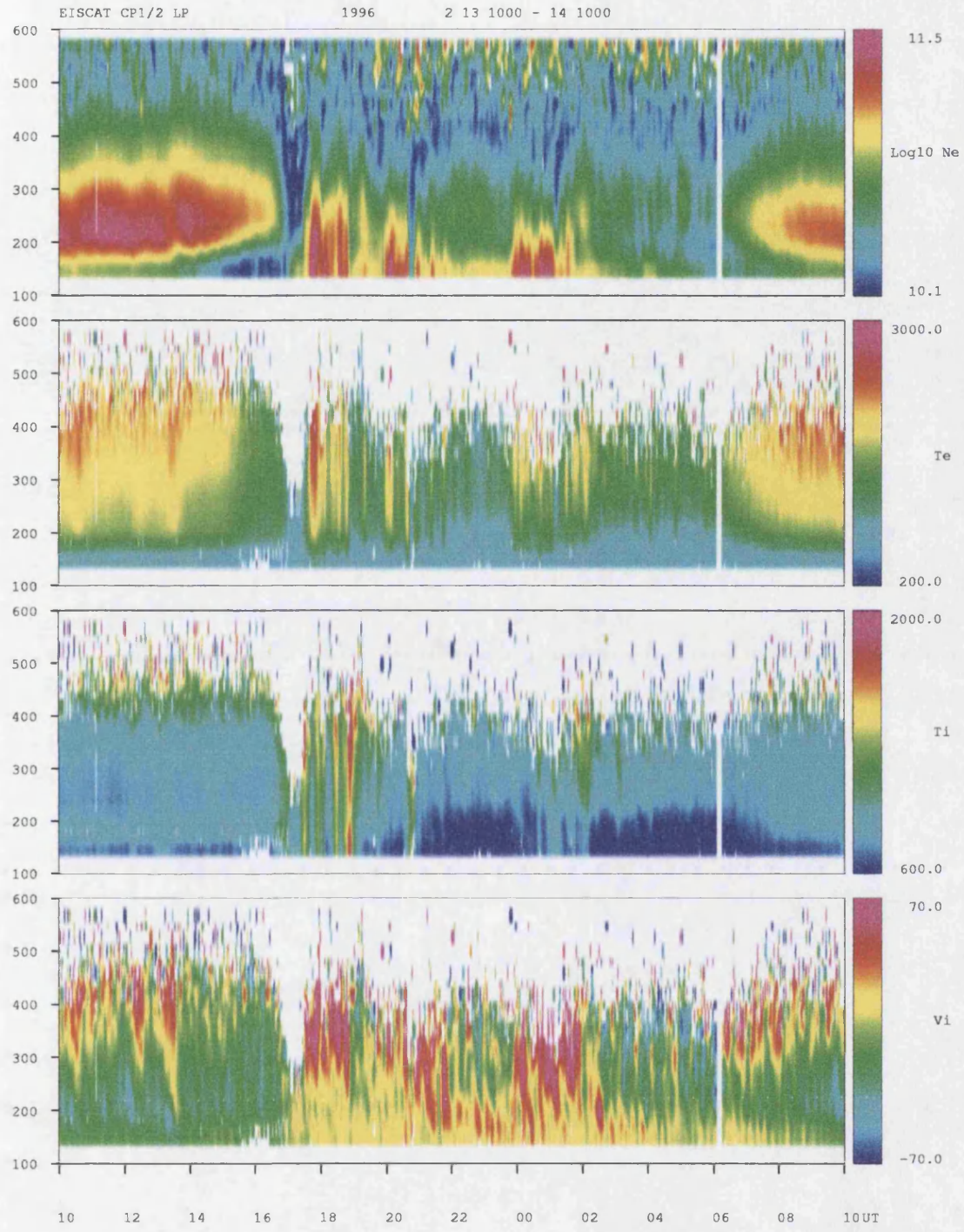


Figure 3.13 Example EISCAT data from 13-14/2/96. From top, plots show Log Electron density, Electron temperature, Ion temperature and Ion velocity.

Of the common programmes CP-5 is rarely used and CP-4 and CP-7 have also been used in dual radar operations (in this mode the designation changes to SP). The CP-1 (UHF) experiment uses a fixed transmitting direction along the local magnetic field to measure electron density, electron and ion temperatures and velocity along the radar beam as functions of altitude. The full, three-dimensional plasma velocity vector and anisotropies in temperature are measured using all three EISCAT sites including the receiving stations at Kiruna and Sodankylä. CP-1 is designed to provide results with very good time resolution and on longer time scales measurements are also used in studies of diurnal, seasonal and solar cycle variations.

Figure 3.13 shows the time series of measurements taken by EISCAT during a run of the CP-1 programme on 13-14/2/96. This represents a "quicklook" of results from long pulse data and shows the electron densities, electron temperatures, ion temperatures and ion velocities. Common Programme data that has been processed and had the relevant parameters extracted is saved in so-called NCAR-format at RAL. Using this database it has been possible to write a routine in the MATLAB application to take NCAR-format result files and analyse the electron density measurements to derive values of *hmF2*, for input to the MWM to produce equivalent neutral winds. This is dealt with in more depth later in this chapter.

The ability of the EISCAT UHF radar to make tristatic ion drift velocity measurements allows electric field data to be derived that can be used to correct the neutral winds produced through the MWM using *hmF2* values. The ion drift velocity component perpendicular to the Earth's magnetic field in the positive magnetic north sense, along with knowledge of the Dip angle *I*, allows the electric field correction to be calculated. Again the methods used are explained later in this chapter.

3.5.3 The Swedish Ionosondes

One of a group of three Swedish Ionosondes is located in Kiruna close to the IRF. This co-location with the Kiruna FPI allows comparisons to be made between the two instrument datasets. Another of the ionosondes is located 360 km south of Kiruna at Lycksele (64.7 N, 18.8 E) and measurements from this instrument could be used for comparison also, given the large horizontal spatial scale lengths for neutral winds compared to the separation of the sites. The third ionosonde is located 880 km south of Kiruna at Uppsala, but none of the data from this site has been considered for comparison

given the separation of the sites. All of the Swedish ionosondes are digisondes similar to that at RAL, and were installed in early 1990. Before this, analogue ionosondes operated at the sites from their installation in the late 1950's. The routinely scaled ionospheric parameters which can be used as input to the MWM have been recorded at these sites throughout their operation and are available in IIWG format at the WDC-C1 storage facility at RAL.

3.5.4 EISCAT Dynasonde

Complementing the EISCAT radar facility at Tromso is the EISCAT Dynasonde which is a sophisticated and powerful modern HF ionospheric sounder. The system at the EISCAT site near Tromso consists of a 10 kW transmitter with a frequency range of 100 kHz to 30 MHz, a log-periodic transmitting antenna, a spaced-array of dipole receiving antennae, a clock, two receivers and a PC computer.

The Dynasonde measures the time-of-flight (TOF), amplitude and phase of ionospheric echoes as a function of frequency and time. For each transmitted pulse set, several ranges may be determined from which echoes with amplitudes greater than a specified minimum are recorded. The phase for each echo range then provides a set of simultaneous equations parametric in frequency, time, receiving antenna position and orientation. These equations may be solved to derive physically-meaningful parameters such as echo-location, group range, Doppler velocity and polarisation.

Data from the Dynasonde is more time consuming to analyse for the purpose of this thesis than that from the Digital Ionosondes, however, as it is not routinely scaled for the necessary parameters used as input to the MWM. It is also not in round-the-clock operation and therefore the number of times coincident useful measurements are made with the Kiruna FPI are limited. Deriving *hmF2* values directly from the recorded ionograms from the Dynasonde is possible using the POLAN routine within an interface developed at RAL, however this is a very labour intensive and subjective process.

3.5.5 Combined Measurements

A particular advantage of the high latitude site is the large existing database of both neutral winds from the FPI at Kiruna, dating back to 1981, and EISCAT common programme data which may be used to derive values of *hmF2*, also dating back to 1981. Additional data are available from both the Kiruna Ionosonde and EISCAT Dynasonde.

This is of great benefit to the present study as it allows a wide range of seasonal, geomagnetic activity and solar cycle conditions to be investigated for effects on the neutral wind derivation process, using the various sources of *hmF2*.

The large horizontal scale length associated with the neutral wind field in the thermosphere allows comparison between the datasets of the instruments even though the sites are separated by 160 km. However the abundance of data from the sites also allows the corroboration of this assumption. The EISCAT data has been selected to fulfil a number of criteria. Only data taken from CP experiments which are directed towards Kiruna (e.g. the field-aligned position from the CP-3 scan) have been considered and also only those experiments with tristatic velocity data, in order to allow the electric field correction to be applied to the derived winds.

Some of the Kiruna ionosonde data will be presented later in this thesis in comparisons covering a few days but unfortunately no large database with significant seasonal or solar activity covreage is available from this, or any of the other Swedish ionosondes for neutral winds as the output generated from the MWM is very limited. This is partly due to the more variable ionospheric conditions at high latitudes producing input which falls outside the expected range for the MWM and thus no winds are generated. The EISCAT Dynasonde has not been used due to the restricted coverage in its database and the effort required to extract *hmF2* values from this database.

3.6 Details of the neutral wind derivation techniques

For each of the implementations of neutral wind derivation applied to the instruments described in this chapter, several practical and instrument specific details need to be presented before results can be placed in proper context. This section will outline the process involved in deriving each set of neutral winds.

3.6.1 Digisonde winds using MWM

A feature of the MWM is its ability to process data in IIWG format to estimate *hmF2* values and then use these values to derive neutral winds. Many ionosondes around the world forward their data in IIWG format to the WDC-C1 at RAL. This database extends back to 1932 for the Slough/RAL ionosonde and has well over 100 instruments currently contributing. In this study data from both the Kiruna Digisonde and RAL Digisonde from the WDC database have been used. In order to use the MWM with this

data the daily mean A_p, sunspot number and $F_{10.7}$ value for the day need to be established.

An alternative to this method would be to use true height analysis on the raw ionograms to establish *hmF2* values. However this technique requires subjective choices to be made in the analysis. Algorithms such as POLAN and ARTIST have been developed to implement true height analysis but results are not qualitatively different from those derived using standard scaled parameters. Even this technique, using these algorithms, is work intensive as automatic scaling from ionograms is estimated to be accurate only 60% of the time and requires checking by experienced scalers. In summary, none of the data presented in this study has been derived using true height analysis.

IIWG format data provides hourly data for the scaled parameters but in the process of conversion to neutral winds some of the data will lie outside the acceptable range for the MWM and no output will be produced. For the Kiruna data in particular there have been a large number of nights where there have not been enough data for useful comparison to the FPI and other neutral wind sources. As mentioned in the previous section, for this reason no climatological data is presented based on the Kiruna Digisonde results.

3.6.2 EISCAT winds using MWM

The data from EISCAT experiments may be analysed to derive values of *hmF2*. In this study data have been used from Common Programme experiments, predominately CP-1, stored in NCAR format on the EISCAT project computers at RAL. Matlab ™ routines were written to take data in this format and produce *hmF2* values by fitting a polynomial to the altitude profile of electron density values. The polynomial is then used to calculate the altitude of maximum electron density, which is used as the derived *hmF2* value. The routine is presented in Appendix B. Values arrived at in this manner are input to the MWM along with the relevant values of the various indices for the date in question to produce output in the form of a time series of meridional neutral wind values.

These values however are uncorrected for the influence of electric fields. NCAR format files containing ion velocity measurements with respect to the geomagnetic field are available from many EISCAT experiments and these were used to establish the correction necessary to be applied to the winds output from the MWM. This is achieved by using the programs developed for the INDI experiment to derive tristatic ion velocity

measurements from the EISCAT experiments. These are then converted to equivalent electric field values.

Buonsanto et al. (1989), using ISR data from Millstone Hill from September 17-24, 1984, derived meridional neutral winds from servo theory and also assessed the electric field correction to be applied to these winds using the same data. This study used the earlier work of Rishbeth (1967) and Rishbeth et al. (1978) to show that this correction may be taken from the following representation of the poleward neutral wind (U_N) :

$$U_N = (v_{perpN} \cos I - W) / (\cos I \sin I) \tag{3.19}$$

where v_{perpN} is the ion drift velocity component perpendicular to the earth's magnetic field in the magnetic north direction, I is the magnetic dip angle and W is the vertical drift in *hmF2* applied by a neutral wind or electric field. From this equation an electric field correction, given by $v_{perpN} / \sin I$, can be applied to the U_N derived from servo theory. The Matlab routine written to take the ion drift velocity data from the CP-1 experiments and calculate this correction for the winds derived here from the *hmF2* measurements is presented in Appendix C.

Figure 3.14 shows both the uncorrected and corrected winds from one day of an experiment on 6/8/85 together with the electric field correction which has been applied. A running average was applied to the electric field correction values, as there is usually a lot of variability associated with the electric field measurements, and hence they appear smoother than the derived winds in this example plot. The large inertia of the neutral atmosphere leads to confidence that this process, which is averaging over some tens of minutes, represents a realistic correction to the derived winds and also avoids introducing a source of greater variability to the final corrected winds. The EISCAT-MWM winds presented later in this thesis have been corrected using this method. Electric field corrections are usually small during the day but at night can be large, indicating greater geomagnetic activity. In Chapter 5 the criteria for rejecting winds from active periods for the data produced by this method will be explained.

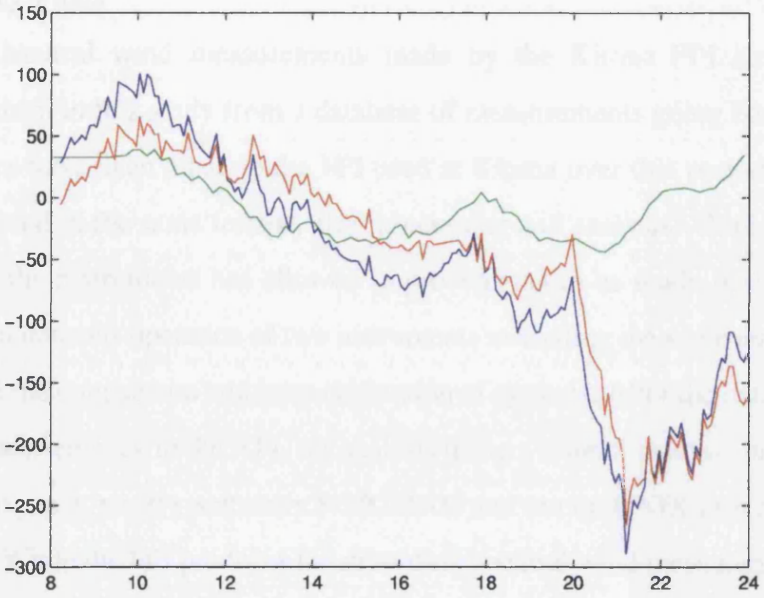


Figure 3.14 Meridional winds derived from EISCAT *hmF2* estimates. The x-axis represents hours UT on 6/8/85 while the y-axis represents positive northward meridional winds in ms^{-1} . The red line represents the uncorrected equivalent winds, the green line represents the electric field correction and the blue line represents the final corrected winds.

3.6.3 EISCAT winds using ISR method

The most commonly used way of deriving neutral winds from incoherent scatter radar data is to use the values of ion velocity parallel to the Earth's magnetic field together with calculations of the diffusion velocity along the field line. This technique has been used with data from incoherent scatter radars at many sites (e.g. Millstone Hill, EISCAT, Saint-Santin, Arecibo) to produce neutral winds from experimental data. An example of the application of this technique is the EISCAT INDI experiment, which was described in Chapter 2. The INDI analysis technique has been used to analyse some of the EISCAT data presented in Chapter 5 for comparisons limited to a few days. The technique has been described previously in Farmer et al. (1990) and Davis et al. (1995).

A number of climatological studies have also used similar techniques to look at average meridional neutral winds from EISCAT data under various conditions (e.g. Titheridge, 1991; Witasse et al., 1998). The comparison of the results from these studies with those presented here will be discussed in Chapter 5. As the Witasse et al. (1998) paper produces a climatology of neutral winds based on the standard ISR technique from the EISCAT data, the INDI technique was not used to produce a similar climatology to avoid duplicating this work.

3.6.4 FPI winds

Neutral wind measurements made by the Kiruna FPI have been used for comparison in this study from a database of measurements going back to 1981. Many upgrades have been made to the FPI used at Kiruna over this period but the data have been stored in the same format, after processing and analysis. Replacing the detectors used in the instruments has allowed improvements to be made in time-resolution and contemporaneous operation of two instruments measuring the same airglow emission has allowed the comparison and inter-calibration of systems. After the data are recorded they are transferred back to the APL for analysis using a suite of analysis programs developed over the years, mostly written in FORTRAN and run on UNIX processors.

While the FPI produces the most direct neutral wind measurements, observations are limited to clear, dark, nighttime conditions. The sensitivity of most detectors used for the observation of airglow with FPI devices is such that they could be destroyed by exposure to light from sunlit skies. The enormous solar continuum contribution at the wavelengths of interest during daytime would also require a more complex spectral filtering regime to be adopted by comparison with "normal" single etalon FPI's.

At Kiruna conditions dictate that no appreciable observations are possible from the end of April until the beginning of September, however in mid-winter data may be recorded on a near 24 hour basis. At a mid-latitude site such as RAL it is possible to obtain data on all clear nights throughout the year, however the extent of coverage expands and contracts to follow the seasons.

3.7 Summary

The details of the instrumentation and techniques used to derive the measurements relevant to this thesis have been presented in this chapter. Emphasis has been placed on explaining the techniques and development involved in the FPI measurements in particular as this work has been completed internally within the APL. The FPI thermospheric wind measurements are also the most direct and are therefore used as the primary experimental data to which the external sources of wind data, derived from ionospheric measurements or modelling, will be compared. Relevant background information has been presented to place these external sources and the combination of the instruments at the separate sites in proper context, including some relevant example results.

Chapter 4 – Comparison of meridional thermospheric neutral wind sources.

4.1 Introduction

Theoretical predictions of thermospheric neutral winds can be tested by looking at long term data trends and sensitivity to different conditions, to establish whether the theoretical results are demonstrating the correct range of variation and behaviour. Satellites such as Dynamics Explorer 2 could provide global verification of circulation patterns for theory, but long-term spatially localised databases are needed to test theoretical climatologies. In these cases sufficient data need to be gathered to allow confidence in the comparisons. One of the aims of this thesis is to use one of the largest local experimental datasets available to test both theoretical and empirical model predictions of the neutral winds at a specific location.

In this chapter relevant published comparisons using similar techniques to derive thermospheric neutral winds are presented and discussed. The most important sources of thermospheric neutral wind data currently available for comparison are identified and presented in a consistent form to allow easy intercomparison for climatological results. These include climatologies based on empirical measurements, from an empirical model and a numerical model derived from first-principles. The development of these individual sources as well as the relevance of their results are discussed. These existing climatologies will be used in Chapter 6 for comparison with those prepared specifically for this thesis and presented in Chapter 5.

4.2 Comparisons limited to a few nights

The large majority of investigations of the behaviour of thermospheric neutral winds are confined to periods of a few consecutive nights or a series of nights sharing the same seasonal, solar activity or geomagnetic activity conditions. As a consequence of the mean lifetime of ionization in the F-region often being several hours, the distribution of ionization at any time is greatly dependent on transport

effects due to electric fields and to neutral winds. In using neutral winds derived from ionospheric parameters to investigate what might be termed the "weather" in the ionosphere and thermosphere it is important to be aware of the influence of these effects. The intercomparison of the derived neutral winds with FPI, empirical model and TGCM model winds is therefore important to our understanding and interpretation of the techniques and databases derived in this manner. A number of studies have been carried out using various combinations of instrumentation and/or theory over limited periods to establish the behaviour of the thermospheric neutral wind at different locations and for a variety of conditions. Examples of this type of work follow.

Comparisons of winds derived from EISCAT CP-1 experiments to HWM87 model winds were presented by Titheridge (1991). Wind variations on individual days in summer are shown in Fig 4.1. The EISCAT data were mostly at 20 minute intervals, and successive values were smoothed with a 1:2:1 weighting to reduce point-to-point fluctuations. In all cases the pre-noon and pre-midnight winds are much smaller than those given by the HWM87 model (shown as a heavy line). The discrepancy is greatest near 0800 local time when small equatorward winds are consistently observed, rather than the strong poleward winds predicted by the model. Titheridge used this comparison to point out the limitations of the model at high latitudes, leading to inclusion of some EISCAT and FPI data in the later HWM90 model. As we will see this improves the HWM performance at high latitude considerably.

More recently Buonsanto et al. (1997a) have compared meridional neutral wind results from four first principle models with results from the Millstone Hill ISR, using the line of sight ion velocities, and Fabry-Perot interferometer measurements taken during January 24-26, 1993, a period which included a minor geomagnetic storm. The models used in this study were CTIM, the FLIP model and the NCAR TIEGCM with and without forcings from the AMIE technique. The winds from CTIM, TIEGCM and those derived from the radar were calculated at a height of 300 km. FPI wind measurements were not available for the night of January 24-25. Figure 4.2 shows the horizontal neutral wind in the magnetic meridian plane (*Umer*), positive poleward, given by the various methods.

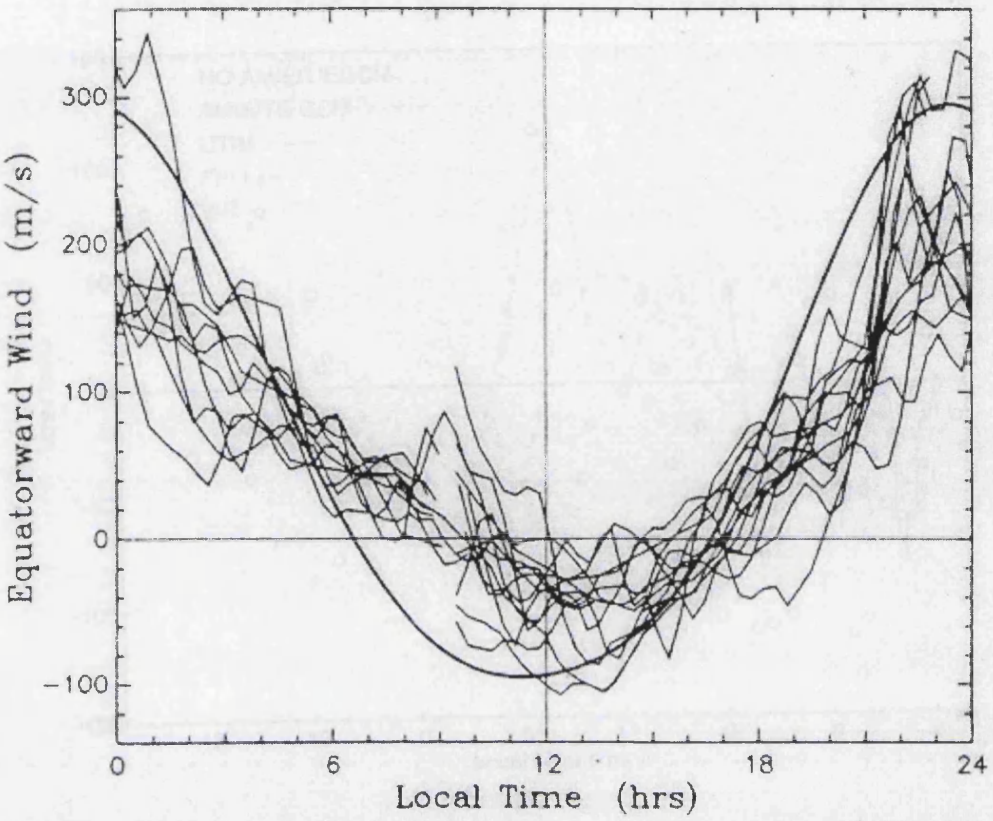


Figure 4.1 Neutral winds derived from low solar activity EISCAT data for individual summer days (continuous lines), compared with the result from the HWM87 model (heavy line) (after Titheridge, 1991).

The winds derived from the radar data are generally more southward than the model winds over the 2-day interval. The authors speculate that as the thermospheric circulation is set up by pressure gradients originating in high latitude Joule and particle heating an underestimate of the high latitude Joule heating could be the cause of this difference between the radar and the winds from these coupled models. Previously, Hagan (1993) calculated average quiet time wind patterns at 300 km altitude from Millstone Hill radar data and found that the meridional winds at solar minimum had a strong equatorward component, in agreement with the radar results shown in Figure 4.2.

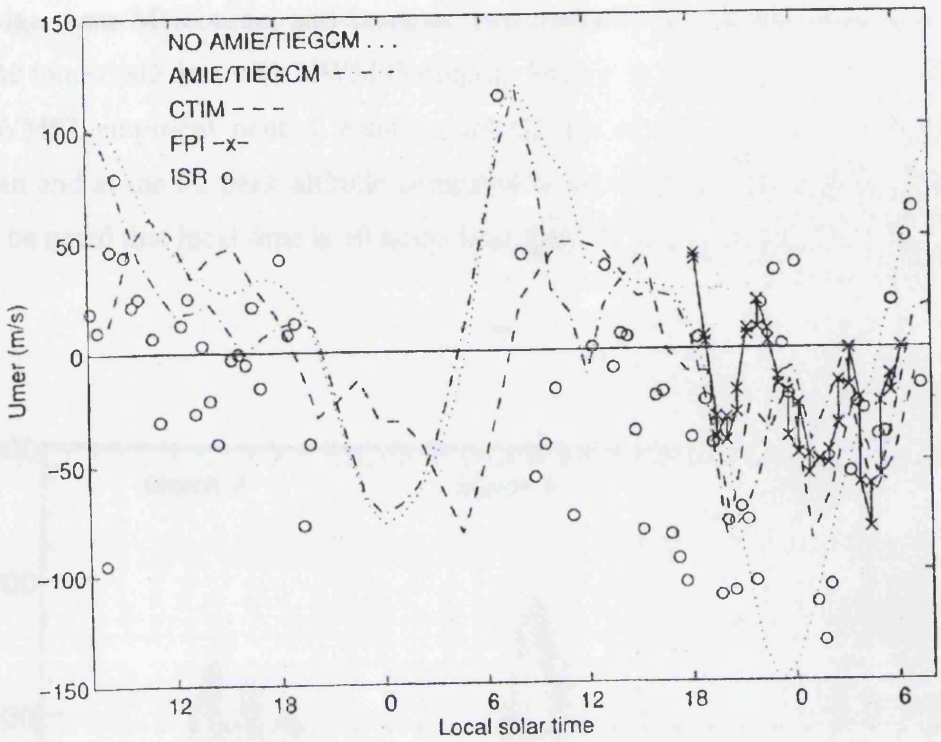


Figure 4.2 Meridional neutral wind (*Umer*), positive northward, on January 24-26, 1993, observed by the Millstone Hill Fabry-Perot interferometer (crosses) and observed at 300 km altitude by the Millstone Hill incoherent scatter radar (circles), compared with *Umer* at 300 km given by CTIM (dashed curve), the AMIE TIEGCM (dot-dashed curve), and the TIEGCM without AMIE input (dotted curve) (after Buonsanto et al., 1997a).

The night of January 25-26 was characterised by recurrent equatorward surges in the neutral wind, followed by abatements. This pattern is seen in both the radar and FPI wind data, and is reproduced by CTIM and the AMIE TIEGCM. The authors postulate that the equatorward surges are associated with the passage of travelling atmospheric disturbances seen in both the AMIE TIEGCM and CTIM results. These surges are not seen in the TIEGCM results without the AMIE input because of the considerable smoothing of the model inputs with this version. These results demonstrate the ability of the TGCM models to reproduce complex patterns seen in the experimental data reasonably well over these short periods when sufficiently complex input data is used. This study will concentrate on the ability to provide longer term climatological data.

Dyson et al. (1997) use FPI and Digisonde derived winds for comparison at Beveridge, near Melbourne, and compare two methods for neutral wind derivation from the ionosonde data with HWM93 output. Shown in Figure 4.3 is the wind from the HWM93 empirical neutral wind model (Hedin et al., 1996) in the magnetic meridian and at the F2 peak altitude compared to winds from the ionosonde data. It should be noted that local time is 10 hours later than UT at this location.

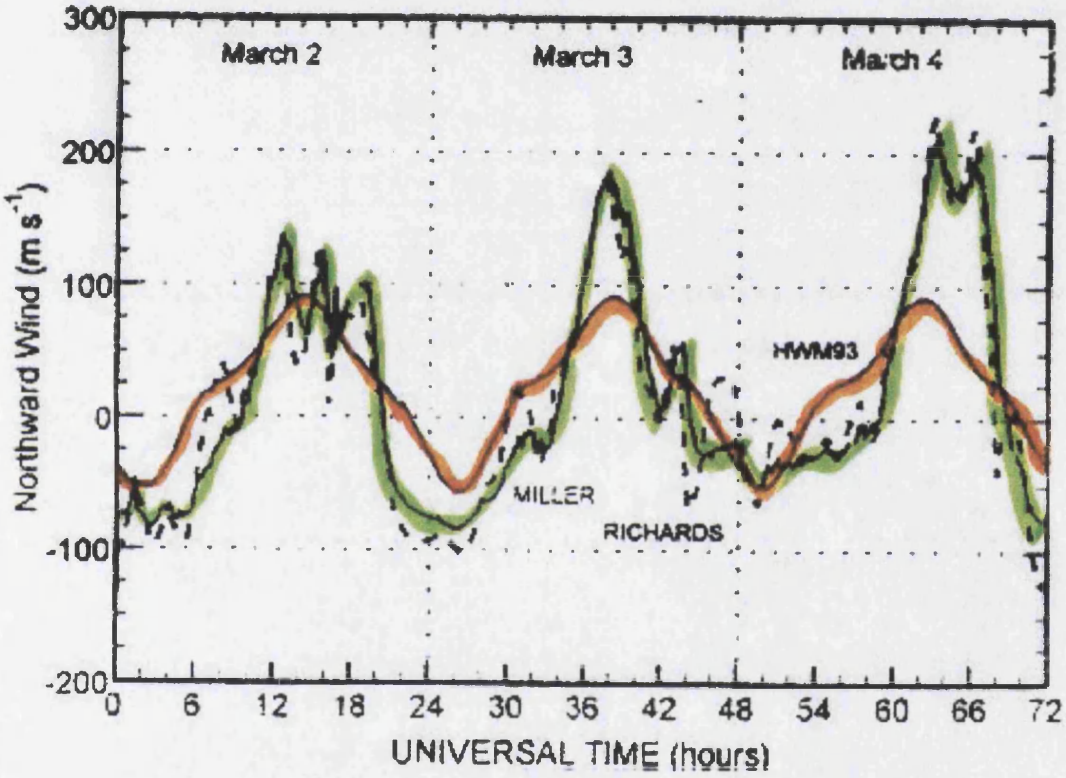


Figure 4.3 Thermospheric meridional neutral winds derived from *hmF2* values by both the Miller et al. (1986) method (solid line highlighted green) and the Richards (1991) method (dashed line) for March 2 to March 4, 1995, at Beveridge, Victoria, Australia (37.5^{o}S, 144.9^{o}E). Also shown are winds from the HWM93 model (red highlighted solid line) (after Dyson et al., 1997).

There is a significant difference in the data used to construct the HWM90 and HWM93 versions of the model, compared to the HWM87 version, in that high latitude ground based input from both FPI and EISCAT data were included. The middle and upper thermospheric winds from the HWM90 and HWM93 versions are identical. The HWM93 model winds agree reasonably well in overall magnitude with

the winds from *hmF2* during the day and on the night on March 2, but are about a factor of 2 smaller than the *hmF2* winds near local midnight on March 3 and 4 which were characterised by more magnetically disturbed conditions.

FPI winds on these nights were in close agreement with the *hmF2* derived winds. Relatively little southern hemisphere data have been included in the HWM93 model so the overall comparison cannot be considered to display a failure of the model by itself, however the poor response to geomagnetic activity is also apparent in the northern hemisphere.

In producing the results for this thesis and preparing the data for climatological work it has been possible to compare sets of individual nights also from the various data sources introduced in Chapter 3. We have seen that the HWM93 version was shown by Dyson to be insensitive to changes in geomagnetic activity and we can trace the same behaviour at this site by looking at the data in conjunction with the geomagnetic activity.

In Figure 4.4 we show HWM93 used in conjunction with Kiruna FPI and Kiruna Digisonde derived winds to demonstrate how the model compares to experimental data at this high latitude site. There is none of the consistent over-estimation of the diurnal amplitude as seen in the HWM87 winds.

In Figure 4.4 we also plot the 3 hourly A_p index for the nights in question to help illustrate the influence of geomagnetic activity. The nighttime period 1800-0600 UT of the 7/8 December is geomagnetically moderate to quiet, with a mean A_p of 8 and the HWM winds fail to match the FPI maximum southward winds. The following night is very quiet with mean A_p of 4 and the HWM winds match the FPI winds to a much closer degree during this night. On the last night of the 9/10 December we see that there are more disturbed conditions with mean A_p of 14 and again the HWM winds fail to match the FPI southward midnight maximum. The daytime northward winds derived from the Digisonde are generally better matched by the HWM winds which is in keeping with the result of the Dyson et al. (1997) study. The HWM modelled winds use the daily mean A_p as an input but, as can be seen in Figure 4.4, the differences in the output winds for different A_p levels are quite small and fail to reflect the large differences seen in the measured winds.

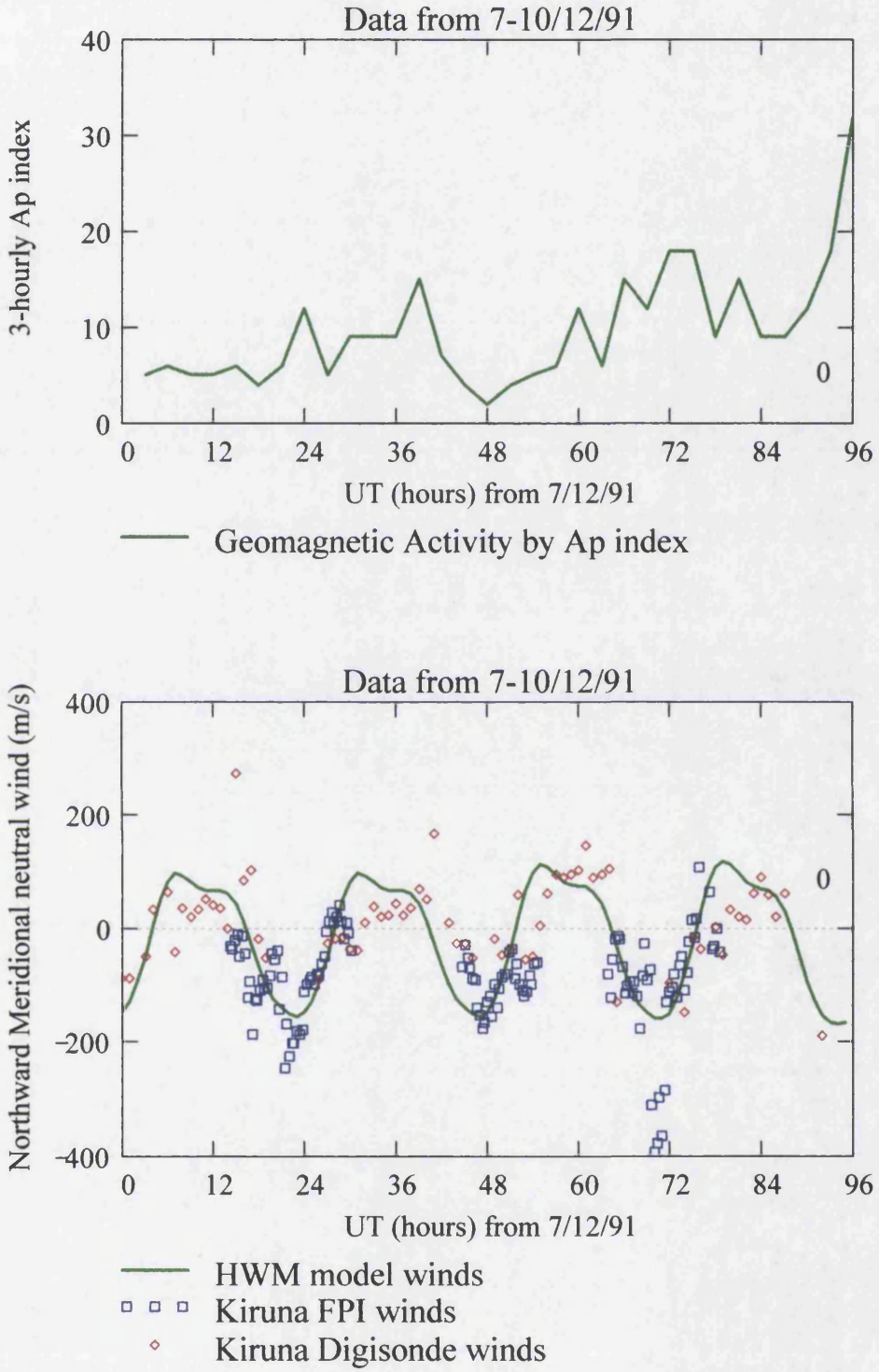


Figure 4.4 HWM93 output compared to Kiruna FPI and Kiruna Digisonde derived winds for varying geomagnetic activity levels on 7-10/12/91.

4.3 Long term climatological comparisons

It is intended in this thesis to concentrate on the comparisons between neutral wind determination techniques based on longer term climatologies rather than the comparisons from individual nights. Hedin et al. (1994) pointed out that identification of a clear dependence of upper thermospheric wind on solar activity has remained an elusive goal, mainly due to a lack of suitable data sources, although Buonsanto (1991) and Hagan (1993) have reported decreasing amplitudes of the diurnal meridional wind with increasing solar activity. While pressure gradients will increase with increasing solar activity and kinematic viscosity decreases, the increased mass of an air parcel and increased ion drag tend to moderate the effect of pressure force increases and may even reduce the wind. Since there may be a rough balance between the factors tending to increase or decrease wind speed with increasing solar activity, these variations can be a sensitive test of theoretical models. This study will allow the effects of solar activity variations, as well as seasonal differences, to be examined from a variety of sources, for a high latitude site, and the implications of the results from each of the sources and their intercomparison will be discussed.

The following sections present the most relevant sources of available data for comparison, in a climatological sense, with relevance to this study for the EISCAT and Kiruna site: from a study by Witasse et al. (1998, hereafter referred to as WLLP), HWM empirical model winds and CTIM theoretical model.

4.3.1 ISR derived winds

An early investigation of the influence of solar activity and season on thermospheric neutral winds derived from line of sight ion velocities was carried out by Duboin and Lafeuille (1992) using data from the Saint Santin, France, ISR. This mid-latitude site had available a 15 year database of winds from 300 km altitude which allowed many of the important features of the data to be identified. They found that the summer winds were systematically more equatorward than the winter winds, a difference which became greater with increasing solar activity and magnetic activity.

The Millstone Hill ISR database, also at mid-latitudes, has been used to produce comparisons for particular intervals (i.e. Buonsanto et al. (1997a)) or to test the influence of changing parameters on the derivation of winds from ionospheric data (i.e. Buonsanto et al. (1997b)) but only limited work has been done on producing a

climatology of neutral winds (Hagan, 1993) until very recently (Buonsanto and Witasse, 1999). The study by WLLP has presented a climatology of neutral winds based on measurements derived from EISCAT experiments using a technique similar to that described as INDI-like in Chapter 3.

Their analysis of sixty-five ISR experiments represents 119 days between January 1984 and March 1995. During most of these experiments the radar was operating in a fixed mode with the Tromso antenna looking along the direction of the magnetic field (CP-1 – like mode). In a few experiments the radar was operating a four-position scan (CP-2 mode) in which case data were selected corresponding to the field-aligned position, thereby extracting CP-1 – like data. As there is a full solar cycle of data, they identify the major strength of their experiment as the ability to study the influence of solar activity on the meridional wind. This aspect was unavailable to Titheridge (1991) given the smaller dataset gathered at the time of his study of the neutral winds derived from EISCAT experiments.

WLLP follow the method described in Lilensten and Lathuillere (1995) to deduce the meridional component of the neutral wind. In this implementation the meridional wind (positive northward) is given by:

$$U_{merid} = \frac{1}{\cos I}\left(-V_{i\parallel} + V_{diff}\sin I - U_z \sin I\right)$$

where I is the dip angle, $V_{i\parallel}$ the ion velocity along the magnetic field line (positive upward), V_{diff} is the ambipolar diffusion velocity, and U_z the vertical wind (positive upward).

I is given by the IGRF model (IAGA working group, 1996) and $V_{i\parallel}$ is directly measured by the radar. V_{diff} is taken from Lilensten and Lathuillere (1995), and it depends on the gradients of the electron density and the ion and electron temperatures measured by the radar. It also depends on the O^+-O collision frequency, through the ambipolar diffusion coefficient. This O^+-O collision frequency is taken from Schunk and Nagy (1978). In the WLLP study this collision frequency has not been multiplied by the Burnside factor (Salah, 1993) as they take into account the conclusions of Lathuillere et al. (1997). These conclusions were reached by comparing

contemporaneous wind profiles above Tromso deduced from EISCAT data and from WINDII interferometer measurements. They found a better agreement between both profiles when this factor was not included in the collision frequency. This agrees with the trend towards a lowering of the Burnside factor towards unity as described earlier in Section 2.6.

The neutral densities are given by the MSIS90 model (Hedin, 1991), for which the exospheric temperature is set to the measured ion temperature at 278 km height (Banks and Kockarts, 1973).

Some data were rejected as, in the presence of a large electric field, the vertical wind can reach values over 100 ms^{-1} in the thermosphere (Smith, 1998). The diurnal, seasonal and solar flux dependence of the vertical wind is not well known although its mean value is generally very small (i.e., Fauliot et al., 1993), which leads most studies to set it to zero. WLLP however chose to keep only the local quiet periods by using the following criteria:

1. When the knowledge of the electric field E is available (three antennas of Tromso, Sodankyla and Kiruna operating simultaneously), $E < 20$ mVm^{-1}.
2. When E is unknown, they rejected data if the ion temperature at 278 km height was larger than the exospheric temperature given by the MSIS model plus 200 K. This is taken as an indicator of Joule heating.

The binning of the data reflected the experience of previous authors and the need for easy comparison with published results. Some of the previous studies split the meridional winds according to two geophysical parameters A_p and $F_{10.7}$ (e.g., Duboin and Lafeuille, 1992), in which $F_{10.7}$ represents the solar activity and A_p represents the geomagnetic activity. In the case of Aruliah et al. (1996a) the K_p index was used instead of A_p. WLLP examined the mean winds at 257 km height for high solar fluxes and two geomagnetic activity levels ($A_p < 11$ and $A_p > 11$). They found no significant difference between these levels, with mean winds, reversal times and midnight winds all very similar. This behaviour was also checked for the low solar activity level. Based on this result they no longer consider A_p as a pertinent criterion and state they believe this is in keeping with their choice to reject data corresponding with local magnetic active periods. It is worth noting that their rejection criteria may

in fact be the cause of the lack of any geomagnetic activity dependence in the data. This issue is discussed later in Chapter 6.

Instead WLLP consider just the two following parameters: season and solar flux. Their sample is divided into two solar activity levels considered active and quiet, with the median value $F_{10.7} = 120$. The seasons have been defined in the following manner: winter corresponds to the period centred around December solstice +/- 45 days, spring to March equinox +/- 45 days, summer to June solstice +/- 45 days and autumn to September equinox +/- 45 days.

The eight classes of wind thus defined contained at least six different days each. WLLP note that the choice of these classes often differs between studies with, for example, Duboin and Lafeuille (1992) considering two solar activity levels, above or below $F_{10.7} = 100$, and Hagan (1993) choosing solar maximum periods associated with $F_{10.7} > 180$ and solar minimum with $F_{10.7} < 90$. The seasons in the Hagan study also differed, containing only two months each. These differences are further discussed in Chapter 6.

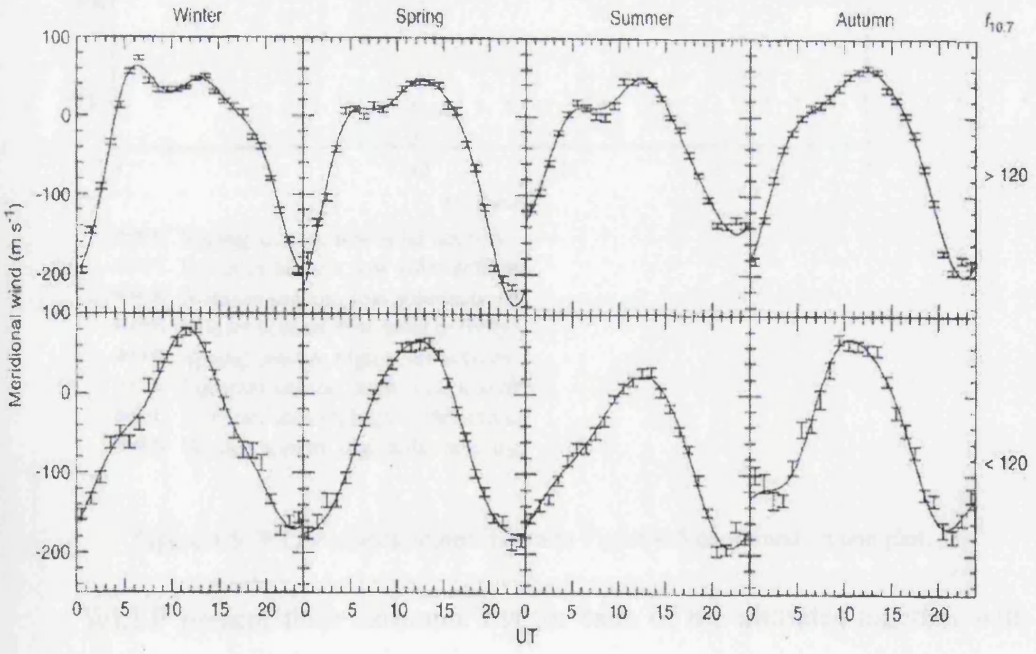


Figure 4.5 Mean meridional wind patterns (points with error bars) and the harmonic fit (solid line) at 257 km height from EISCAT experiments. Seasons are classified as outlined in the text. The upper line of boxes corresponds to high solar fluxes, the lower to low solar fluxes. All data from geomagnetically quiet periods (after WLLP).

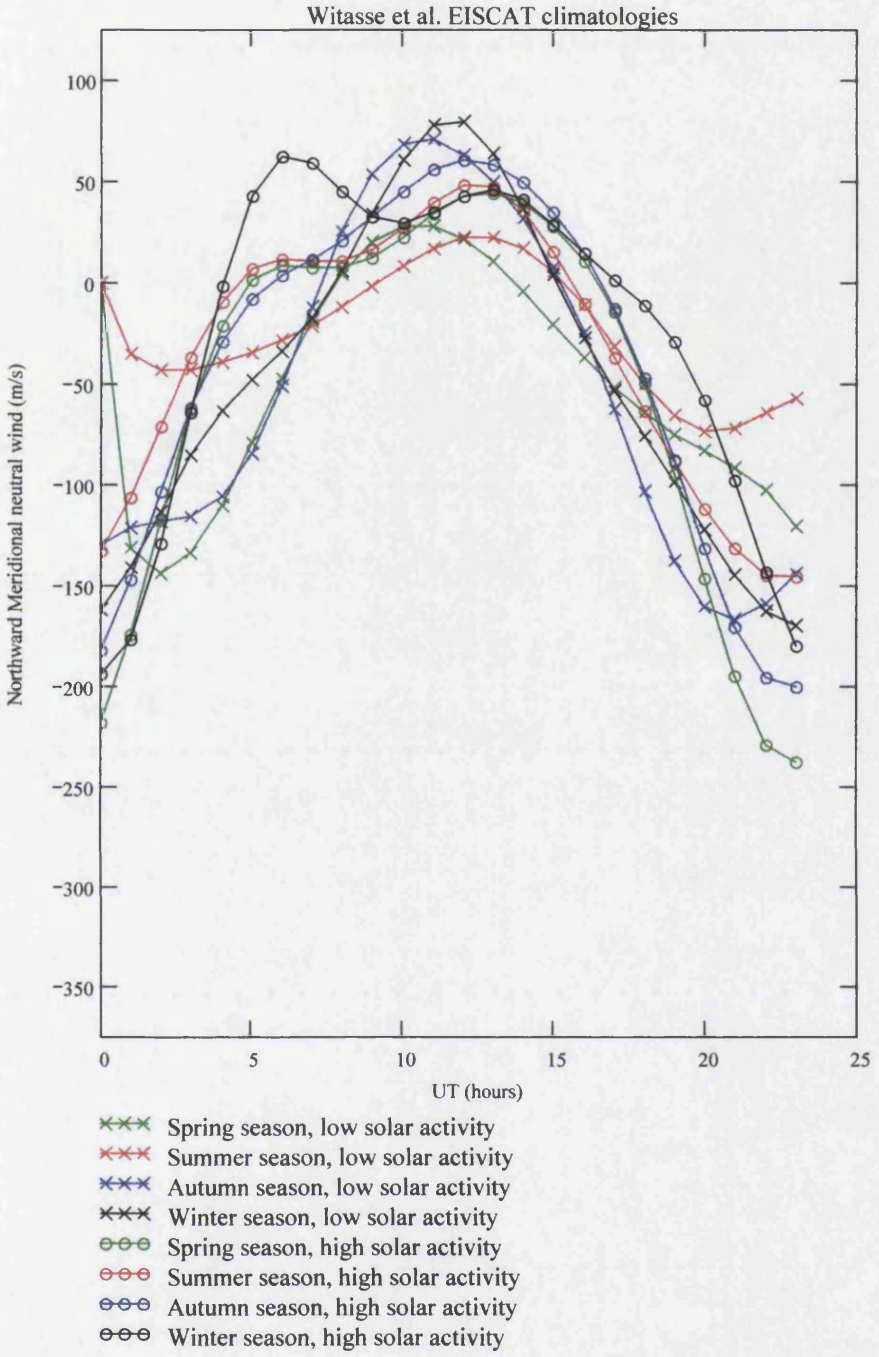


Figure 4.6 WLLP climatologies shown in Figure 4.5 combined on one plot.

WLLP present their harmonic fits for each of the altitudes together with the standard deviations on these fits. For the winds at 257 km, chosen for comparison in this thesis, the standard deviations on the derived winds range in value from 34 ms^{-1} in winter season at high solar activity to 66 ms^{-1} in spring season, also at high solar

activity, with the other regimes having intermediate standard deviation values. The data and their harmonic fits at this altitude are shown in Figure 4.5.

Figure 4.6 shows all of the separate climatologies together in one plot and we see the difference in diurnal variations between the seasons and the solar activity levels. The predominant difference between high and low solar activity plots is that at high activity there are consistently higher amplitudes in equatorward winds around midnight than are seen in the equivalent seasons for the low activity situation. The high solar activity seasons also show enhanced poleward winds compared to the low solar activity winds in the midnight to 0800 UT period. WLLP note that there is a much higher semidiurnal component in the high solar activity winds than is the case for the low solar activity regime. They do not attribute this to any particular source, while noting earlier studies using harmonic analysis at mid-latitude that have attributed the component to a semidiurnal tide propagating up from the lower atmosphere (Buonsanto, 1991; Hagan, 1993). A more detailed discussion of these and other features of the WLLP winds, in conjunction with the results from the other sources used in this study, are presented in Chapter 6.

4.3.2 Horizontal Wind Model output

Over recent years the source most often used for model thermospheric winds when measurements are unavailable has been the Horizontal Wind Model (HWM), introduced in Chapter 2. This model has been developed from its original version HWM87 (Hedin et al., 1988) through HWM90 (Hedin et al., 1994) and HWM93 (Hedin et al., 1996), giving progressively increased coverage of conditions and altitudes. We have seen that Titheridge (1991) used EISCAT data from low activity conditions only, to derive a climatology of meridional neutral winds and to compare to HWM winds. In this comparison the older HWM87 version was used which has no solar activity dependence, but as the data were restricted to periods of low solar activity, this variation could not be addressed with the dataset anyhow.

The meridional winds used in this early version of the model are based on FPI measurements made with the polar orbiting satellite DE-2, mostly at heights of 300-500km. Global variations are represented by an expansion in vector spherical harmonics, with each expansion coefficient including annual, 24 hr, 12 hr and 8 hr components. Results calculated from this model by Titheridge are shown in the lower section of Figure 4.7, for conditions corresponding to the EISCAT data and using

mean A_p values for each of the seasons used in the EISCAT dataset. Note that the scale used here is half that for the upper section of Figure 4.7, showing that the model winds are much larger and change twice as fast as the observed winds in the morning and evening.

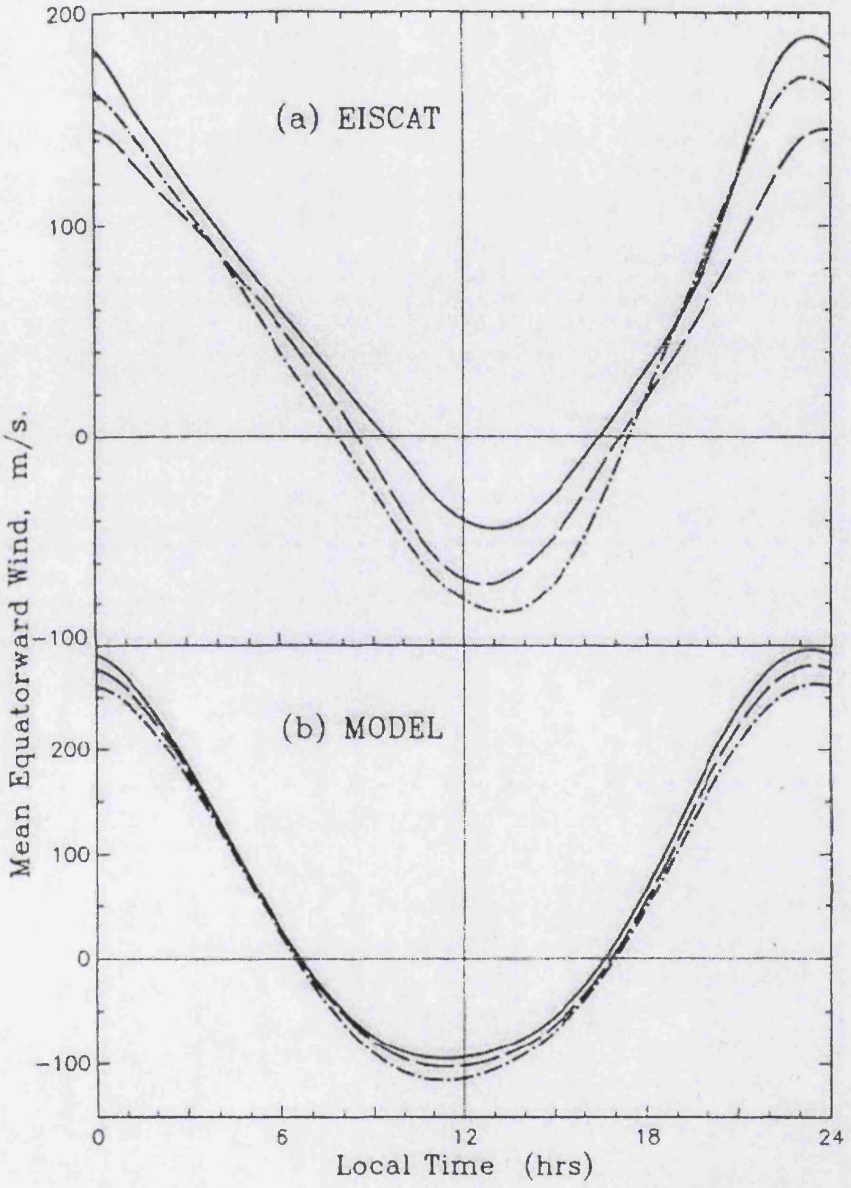


Figure 4.7 Mean equatorward winds for summer (continuous lines), equinox (broken lines) and winter (chain lines) calculated from (a) EISCAT data for 1984-1987 and (b) from HWM87. Note the half-size scale for (b), (after Titheridge, 1991).

There is also a slight difference in phase noted by Titheridge, with the experimental and model variations both having maximum equatorward winds just

before midnight, but maximum poleward winds occur at 11:30 local time for the model and about 13:00 for the observed winds from the EISCAT data. Titheridge performed a Fourier analysis on the curves in Figure 4.7 and the results showed that the overall mean winds in HWM87 are about 50% greater than those calculated from the EISCAT data. The amplitude of the diurnal variation was nearly 100% too large in the model.

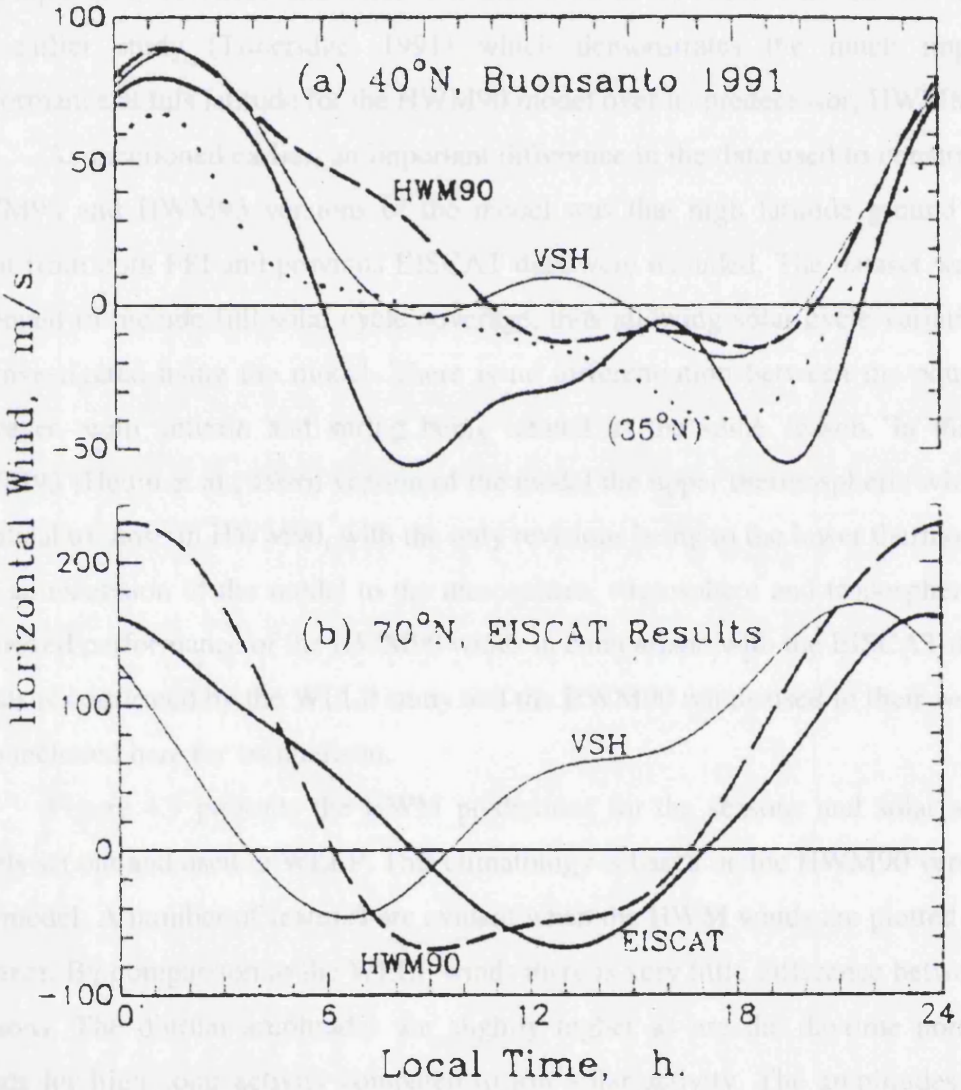


Figure 4.8 Thermospheric meridional neutral winds averaged over summer, equinox and winter from (a) *hmF2* measurements at 40°N (solid line; Buonsanto, 1991) and 35°N (dotted line) (b) ion velocity measurements at EISCAT, 70°N (solid line; Titheridge, 1991). Broken lines in both are HWM90 winds, fine lines are VSH model winds (after Titheridge, 1995b).

Later however, Titheridge (1995b) in a review of empirical wind models finds that for northern hemisphere studies the HWM90 version of the model appears to give the best estimate of mean winds of the available models. For the southern hemisphere, best results were generally obtained by using a mean of the theoretical VSH model and the HWM90 winds. Figure 4.8 shows comparisons of both (a) mid-latitude and (b) high latitude winds derived from ionospheric parameters and their comparison to the output from HWM90 and VSH. The data used in (b) is the same data as used in the earlier study (Titheridge, 1991) which demonstrates the much improved performance at this latitude for the HWM90 model over its predecessor, HWM87.

As mentioned earlier, an important difference in the data used to construct the HWM90 and HWM93 versions of the model was that high latitude ground based input from both FPI and previous EISCAT data were included. The dataset was also extended to include full solar cycle coverage, thus allowing solar cycle variations to be investigated using the model. There is no differentiation between the equinoxes however, with autumn and spring being treated as the same season. In the later HWM93 (Hedin et al., 1996) version of the model the upper thermospheric winds are identical to those in HWM90, with the only revisions being to the lower thermosphere and an extension of the model to the mesosphere, stratosphere and troposphere. The improved performance of the HWM90 winds in comparison with the EISCAT derived winds is confirmed by the WLLP study and the HWM90 winds used in their work are also included here for comparison.

Figure 4.9 presents the HWM predictions for the seasons and solar activity levels set out and used in WLLP. This climatology is based on the HWM90 version of the model. A number of features are evident when the HWM winds are plotted in this manner. By comparison to the WLLP winds there is very little difference between the seasons. The diurnal amplitudes are slightly higher as are the daytime northward winds for high solar activity compared to low solar activity. The amplitudes of the equatorward winds near midnight appear to be consistently slightly higher for the high solar activity conditions than for their low solar activity equivalents. The equinox winds are plotted using the same trace as there is no differentiation between the equinox seasons in the model. Again a fuller discussion of the HWM winds is presented in Chapter 6 when the comparison with other sources is made.

Titheridge (1995b) suggests that "worst case" scenarios in comparison of HWM and data derived from ionospheric parameters give an indication that HWM90

winds may have errors of 50-100 ms^{-1} at some times when the model is used to extrapolate to regions with little data coverage. The winds for Kiruna may be expected to have smaller errors than this however, due to the proximity of a number of the sites which contributed to the overall HWM dataset.

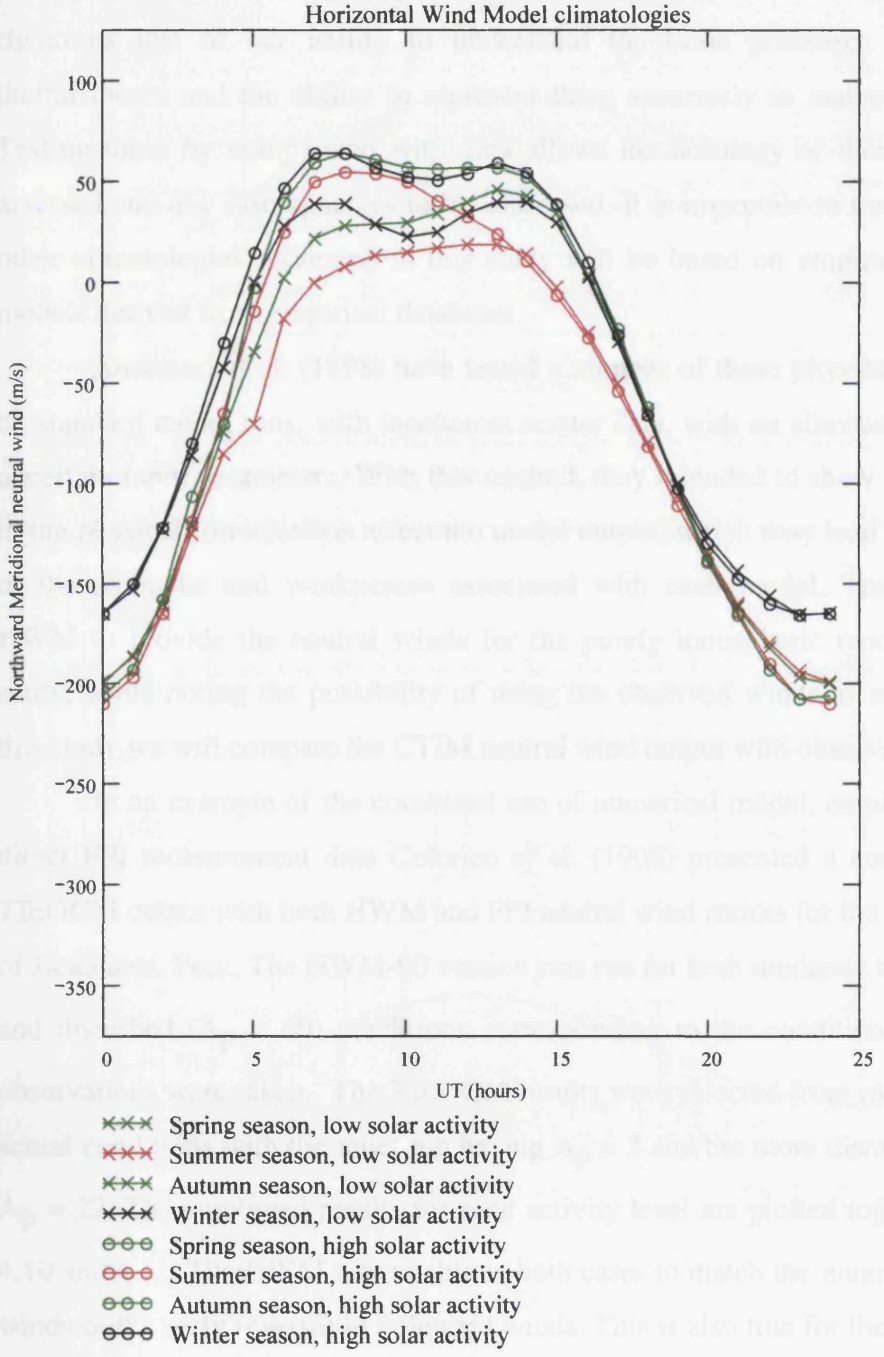


Figure 4.9 HWM climatologies for the seasons and solar activity levels set out in WLLP

4.3.3 CTIM output

In this section the relevant meridional neutral wind results from a standard run of the CTIM model will be presented. In Chapter 2 the development of thermospheric general circulation models which self-consistently solve for the basic equations of state and motion in the coupled thermosphere-ionosphere was presented. These first principles models are very important tools in thermospheric research as they are a rigourous test of our ability to understand the basic processes at play in the thermosphere and the ability to represent them accurately in mathematical models. Testing these by comparison with data allows the accuracy of these models to be assessed and any discrepancies to be addressed. It is important to note that all of the other climatologies presented in this study will be based on empirical databases or models derived from empirical databases.

Anderson et al. (1998) have tested a number of these physical models, based on standard model runs, with incoherent scatter data, with no allowance for adjusting uncertain input parameters. With this method, they intended to show how differences in the physical formulations affect the model output, which may lead to identification of the strengths and weaknesses associated with each model. They however use HWM to provide the neutral winds for the purely ionospheric models used in the study, while noting the possibility of using the observed winds as an alternative. In this study we will compare the CTIM neutral wind output with observed winds.

In an example of the combined use of numerical model, empirical model and direct FPI measurement data Colerico et al. (1996) presented a comparison of the TIEGCM output with both HWM and FPI neutral wind results for the low latitude site of Jicamarca, Peru. The HWM-90 version was run for both moderate to quiet ($A_p = 9$) and disturbed ($A_p = 49$) conditions corresponding to the conditions when the FPI observations were taken. The TIEGCM results were selected from runs closest to the actual conditions with the quiet run having $A_p = 3$ and the more disturbed run having $A_p = 22$. The combined results for each activity level are plotted together in Figures 4.10 and 4.11. The HWM was unable in both cases to match the magnitude of the FPI winds or the early reversal to poleward winds. This is also true for the TIEGCM, with the TIEGCM winds also displaced in time by ~ 2 hours with respect to the HWM90 and FPI results.

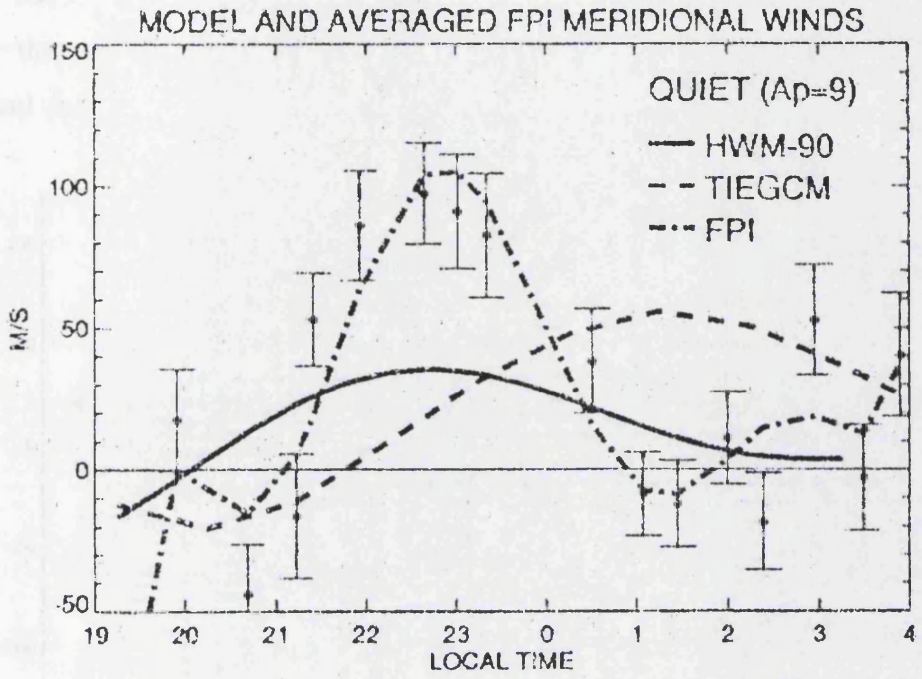


Figure 4.10 – HWM90 model meridional winds for quiet (Ap=9) run at Jicamarca. Also shown are averaged FPI meridional winds and TIEGCM results. Model output reflects conditions on Sept 30, Oct 1,2,8,9 when FPI measurements were taken (after Colerico et al., 1996).

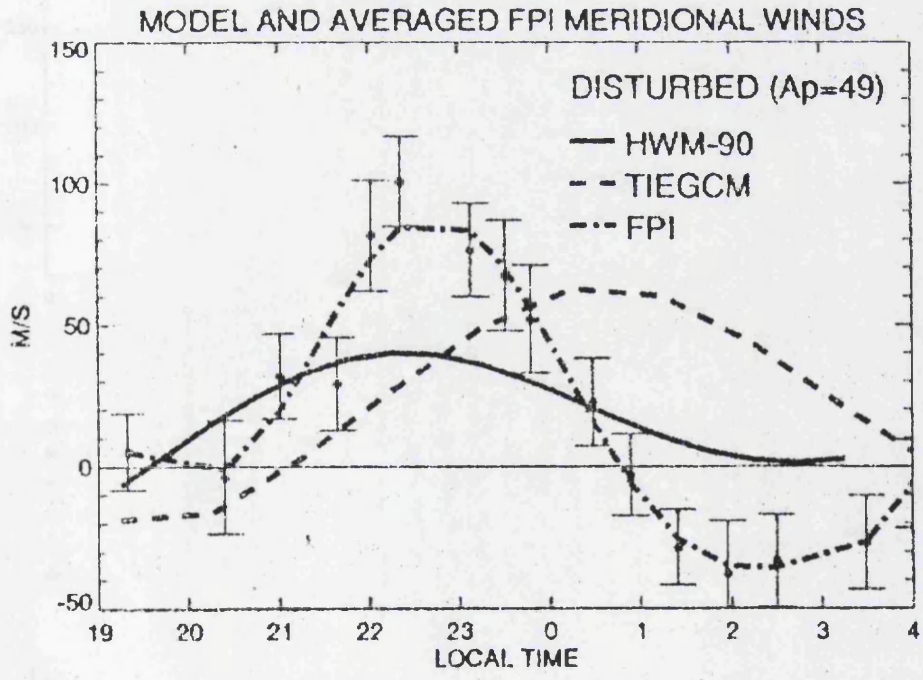


Figure 4.11 HWM90 model meridional winds for disturbed (Ap=49) run at Jicamarca. Also shown are averaged FPI meridional winds and TIEGCM results. Model output reflects conditions on Oct 3-6 (after Colerico et al., 1996).

To examine the ability of the CTIM to produce accurate meridional neutral winds, the following climatology has been produced for comparison with the empirical datasets.

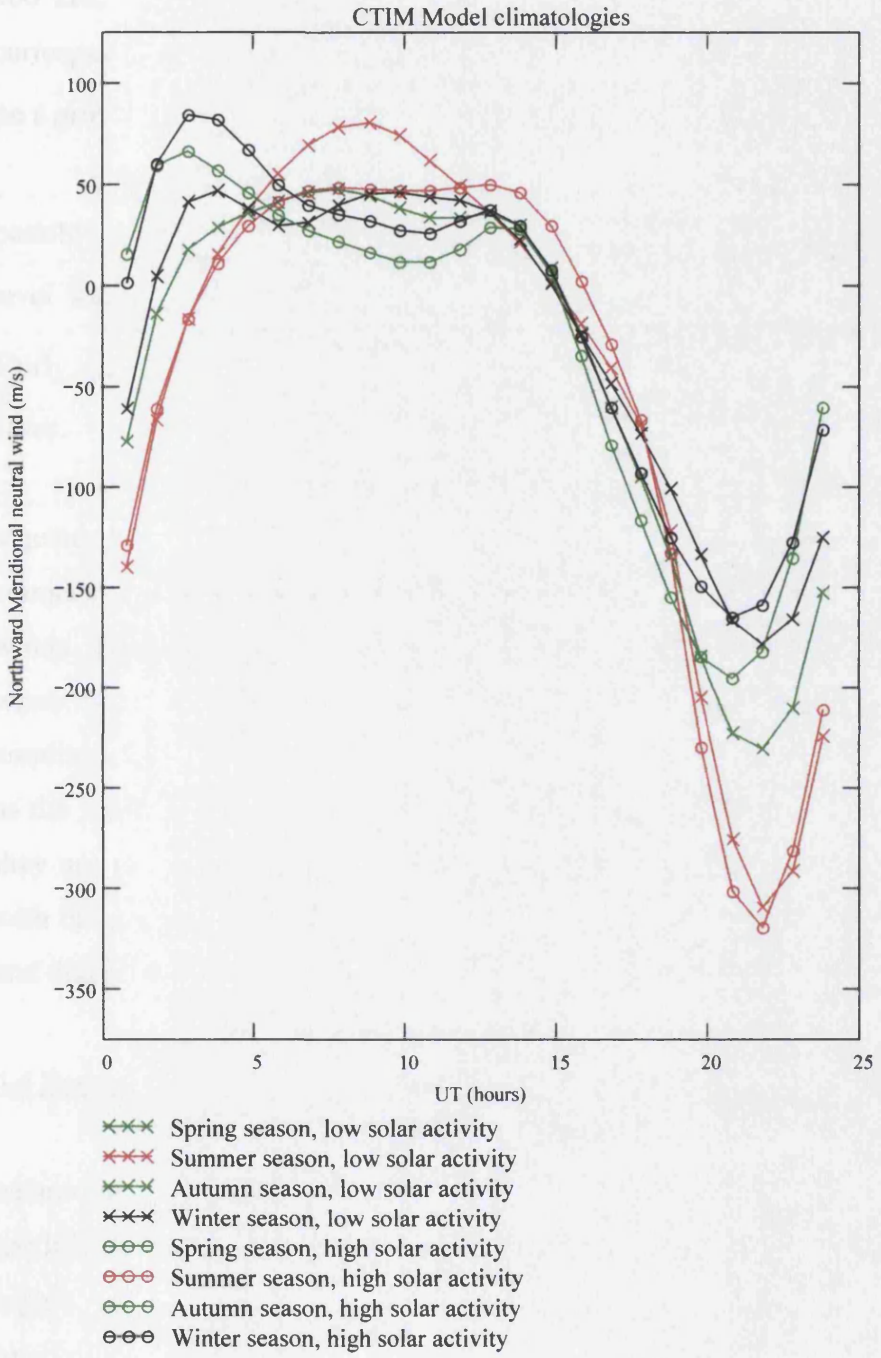


Figure 4.12 The meridional neutral winds at around 240 km from CTIM for 70^{O}N and 18^{O}E for the seasons and solar activity levels set out in WLLP.

The output used for these winds has been obtained from standard runs of the model in which parameters are computed at 15 pressure levels spaced at vertical intervals of one scale height. This corresponds to altitudes from around 80 km to over 400 km, with the wind results presented here being taken from pressure level 12 corresponding to an altitude of around 240 km. The model output is available globally on a grid with spacing 2^{o} in latitude and 18^{o} in longitude.

For this study co-ordinates of 70^{o}N and 18^{o}E were used to match as closely as possible with the EISCAT site. In the case of the low solar activity results the $F_{10.7}$ level was set at 100, with the high solar activity runs having an $F_{10.7}$ level of 180. Fairly quiet geomagnetic conditions with $K_p = 2^+$ were used in both solar activity cases.

The winds extracted from the model runs for the relevant seasons are plotted together in Figure 4.12 where there is a notable consistency to the diurnal variations compared to the WLLP winds although of a different nature to that seen in the HWM winds. The biggest variations are evident in the hours before and just after midnight when the individual climatologies are differentiated by maximum southward amplitudes and morning reversal times. The autumn and spring equinoxes are treated as the same season in this model, in much the same way as with HWM90, therefore they are plotted with the same traces in both solar activity cases in Figure 4.12. As with the previous two sources introduced in this section these winds will be combined and discussed in full in Chapter 6.

4.4 Summary

This chapter has presented the results of some previous studies that have used techniques and sources relevant to this thesis to produce or derive thermospheric meridional neutral winds. These have included both work using sets of individual nights and some longer term studies, all demonstrating the development of these intercomparisons. They have also been used to demonstrate the importance of the use of these sources in combination to both test the models and understand the processes responsible for features seen in the observed data. Three sources have been chosen for comparisons in this study and they represent examples of an empirical dataset, a model based on a variety of empirical input and a first principles TGCM numerical model. The background to their development has been discussed and the relevant

results from these sources have been presented in a consistent format. They will be combined with further results from Chapter 5 and discussed in full in Chapter 6.

Chapter 5 – Climatologies prepared for this study

5.1 Introduction

In Chapter 3 the range of useful instrumentation in the EISCAT – Kiruna vicinity was described and discussed. In this chapter results will be presented from the application of the experimental techniques described in Chapter 3 both to individual sets of nights and also in climatological form. In the previous chapter the development of the most important sources of neutral wind data available in the published research literature, and relevant to this thesis, were identified and some relevant results were examined. An important aim of this thesis is to use one of the largest local experimental datasets of combined thermospheric and ionospheric parameters available to test the neutral winds output by both theoretical and empirical models at a specific location. The most important data sources of thermospheric neutral winds available at Kiruna/EISCAT to be used for comparison are therefore identified and presented in this chapter.

The fact that both the Kiruna FPI and EISCAT ISR experiments have been in operation for over a solar cycle has allowed investigation of both datasets for solar activity variations. Such investigations have been undertaken with the Kiruna FPI database of direct neutral wind measurements (Aruliah et al. 1996a), although the nature of the operation of the FPI instrument means there is less diurnal and seasonal coverage. In Chapter 4 we saw that WLLP had presented a full set of climatologies for season and solar activity level from their EISCAT derived neutral winds. It was also pointed out that only long-term localised databases can test theoretical predictions of long term variations under different conditions. Here we present the wind data from the Kiruna FPI and EISCAT data, derived from *hmF2* values, in the same WLLP format to allow comparison with the previous sources. Also included is a climatology created from *hmF2* values, but in this instance using the *hmF2* output from the empirical IRI model.

These climatologies have been developed specifically for this study and constitute climatologies based on (a) winds from direct measurements, (b) inferred from an interpretation of measured ionospheric parameters and (c) the same interpretation but used with output from an empirical ionospheric model. The previous development and use of the individual sources as well as the relevance of their results are discussed. The range of different techniques is important as it allows investigation of the different approaches and these climatologies will be discussed and the important features arising from this study identified in Chapter 6 in combination with those presented in Chapter 4.

5.2 Individual nights - measurements

In Chapter 3 a variety of experimental techniques were discussed which can be used to produce thermospheric meridional neutral winds measurements. Many of the instruments described in Chapter 3 are located in the Tromso/Kiruna region and so the results from these instruments and their techniques can be compared. Similar, though limited, comparisons have been undertaken in the past by various authors.

We have seen in the previous chapter that studies by Buonsanto et al. (1997a), Dyson et al. (1997) and others have used a variety of these techniques to compare with each other and with model results over short periods. Figure 5.1 is taken from Buonsanto (1991) and demonstrates the comparison of neutral winds derived by two separate techniques based on data from the same Millstone Hill ISR experiment. In this example the winds are calculated from both the line of sight ion velocities at a height of 350 km, similar to the INDI technique in Chapter 3 used with EISCAT ISR data, and from the measured *hmF2* heights, in a method based on the servo model and similar to Miller et al. (1997). No electric field correction has been made to the winds derived from *hmF2* measurements, but the author claims the low geomagnetic activity at the time of the experiment means there is little correction to be made at this mid-latitude site. The major difference between the winds was identified as the difference in heights at which the winds were obtained, with one height fixed, for the ion velocity determinations, and the other varying, from the *hmF2* measurements.

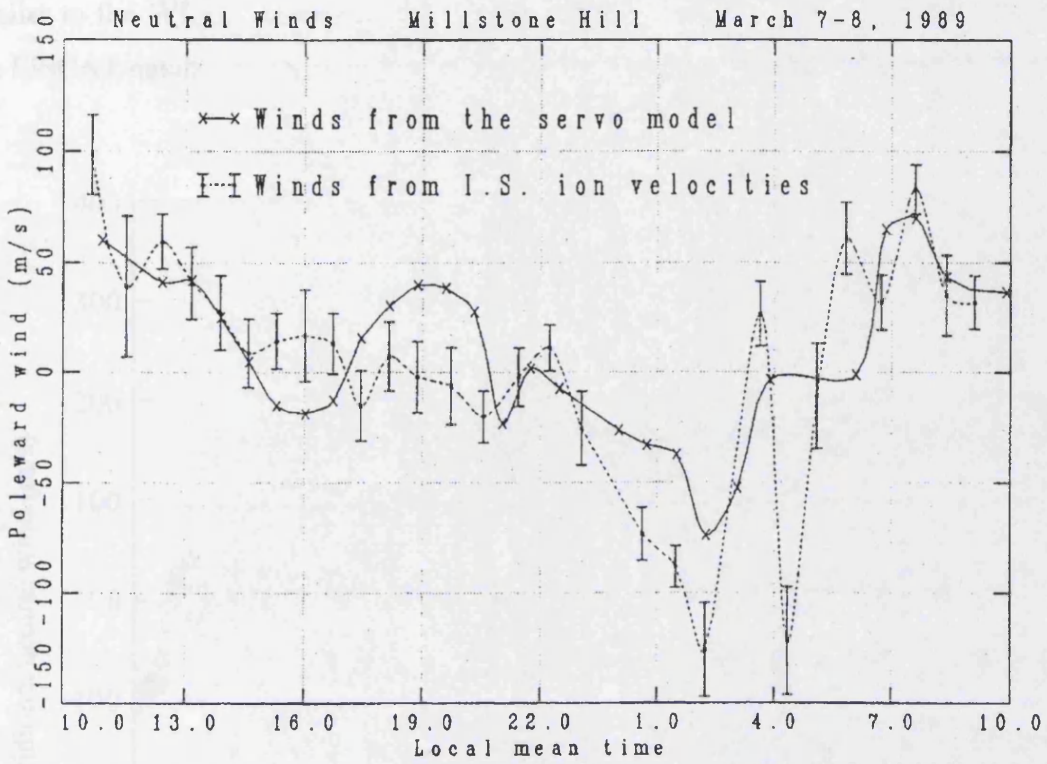


Figure 5.1 Neutral winds at Millstone Hill for March 7-8, 1989. Dots with error bars: winds obtained from incoherent scatter measurements of line of sight ion velocities and calculation of the diffusion velocity at 350 km; crosses: winds obtained from the servo model at the *hmf2* values measured by the incoherent scatter radar (after Buonsanto, 1991).

In this section we present measurements from a series of consecutive days as examples of the data that have contributed to the climatologies based on the various different techniques and shown later in this chapter. These include FPI and EISCAT *hmF2* derived measurements, both of which have been used for these climatologies, as well as Kiruna Digisonde *hmF2* measurements for which only data from individual nights have been derived for comparison. We don't present a climatology based on this *hmF2* data due to a combination of unavailability of a proper source of electric field correction values and a paucity of data relative to the periods covered by the Kiruna FPI and EISCAT ISR.

Also presented are examples of neutral winds derived using EISCAT data with the INDI technique. These are not identical to the winds used for the climatology developed by WLLP, introduced in Chapter 4, since the techniques are independently formulated and implemented. However the INDI winds derived this way will be very

similar to the WLLP winds since the principles are the same, and both methods use the EISCAT database.

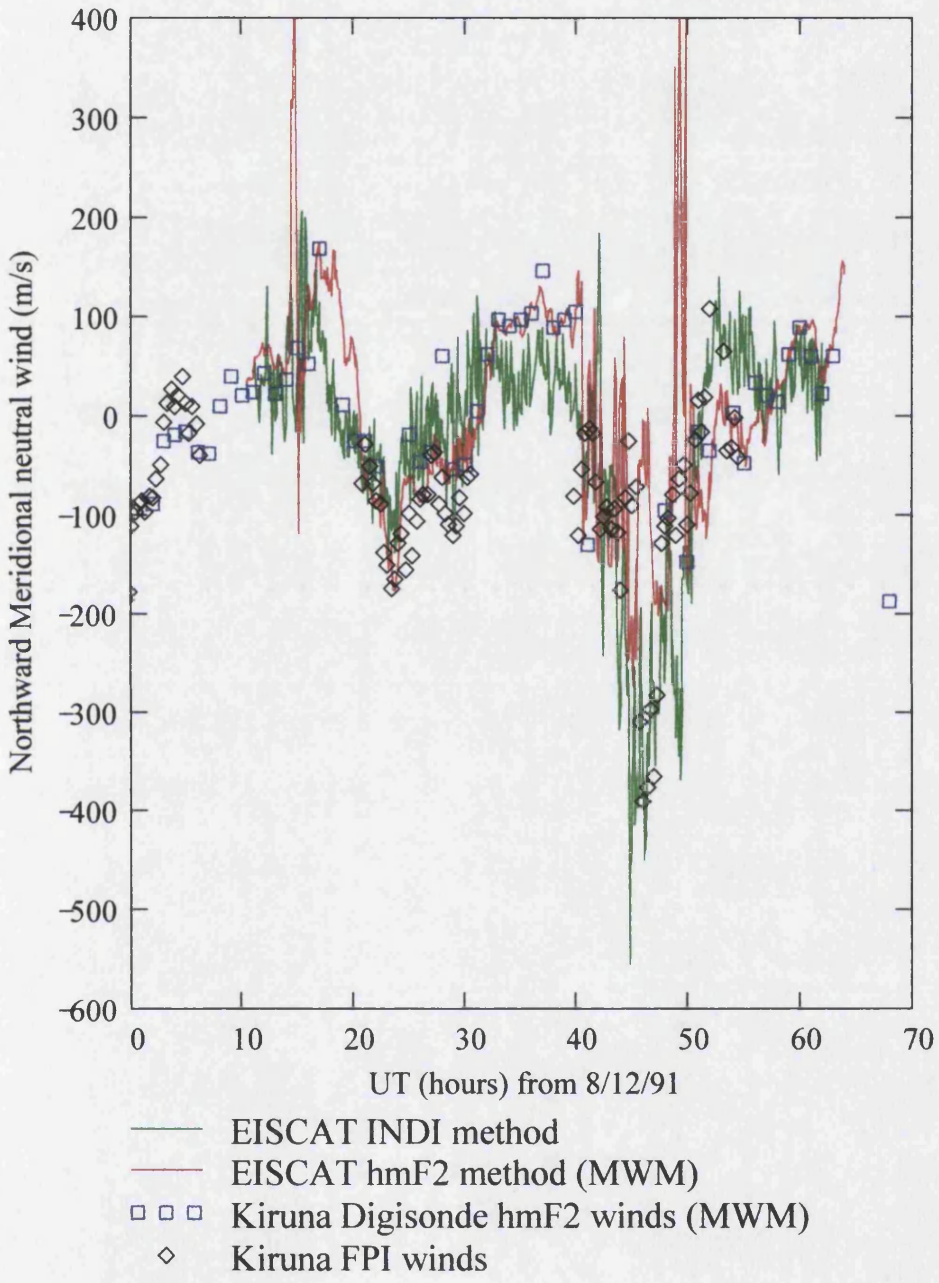


Figure 5.2 Data from three consecutive days beginning at 0000 UT on 8/12/91, showing thermospheric meridional neutral winds from four techniques: Kiruna Digisonde *hmF2* values used with MWM, EISCAT *hmF2* values used with MWM, EISCAT data used with the INDI analysis and Kiruna FPI winds. The consecutive average A_p values for these days were 8,10 and 15. These data contribute to the climatology for winter at solar maximum.

WLLP rely on the technique developed and reported by and Lilensten and Lathuillere (1995) which in turn is based on the earlier work of Salah and Holt (1974). The basis of most studies of neutral winds derived from ISR measurements can be traced back to this technique and these include the techniques used by both Titheridge (1991) and Davis et al. (1995), the latter being the technique used to produce the winds described as EISCAT INDI and presented here.

Figure 5.2 shows the results from a series of days in December 1991. These data have contributed to the winter season, high solar activity climatologies, following the classification scheme invoked by WLLP. The main characteristics in wind patterns at this latitude are demonstrated in Figure 5.2, with northward winds by day and reverse flows at night leading to southward winds. These are the observations of a single site rotating with the Earth under an anti-sunward flow driven by pressure gradients caused by solar heating and high latitude auroral heating. The southward velocities are higher than the daytime velocities due to the lower ion drag encountered after sunset in the absence of photo-ionization, with consequently lower ion densities. Figure 5.3 shows results from a series of days in late March and early April, 1992 which contribute to the spring season, high solar activity climatologies.

Characteristics of the individual techniques are demonstrated on these example plots. The FPI can observe the weak airglow emission above the background only during the dark, nighttime conditions. The Kiruna Digisonde produces derived neutral wind data best during the day as the electron densities used to produce the scaled parameters input to the MWM are usually higher, and consequently more accurately observed, than at night. The EISCAT data is usually available through the entire course of an experiment. However, as with the Digisonde, the accuracy of the results may be expected to drop during the night due to lower electron densities and therefore poor signal to noise ratio which produce less consistent estimates of *hmF2*. The NCAR format result files from which results have been derived usually have constant time resolution throughout the experiment.

In Figure 5.1 we saw a similar comparison for the winds derived from an INDI like technique with a servo type technique for consecutive days at mid-latitudes (Buonsanto, 1991). In these comparisons however there were no coincident FPI or Digisonde measurements available to test the techniques against either direct measurements or the application of servo theory using different instrumentation. The unique database of values from the instruments in the Kiruna-Tromso region allows a

wider range of inter-comparisons for meridional thermospheric wind techniques to be carried out than is possible for any other high latitude site.

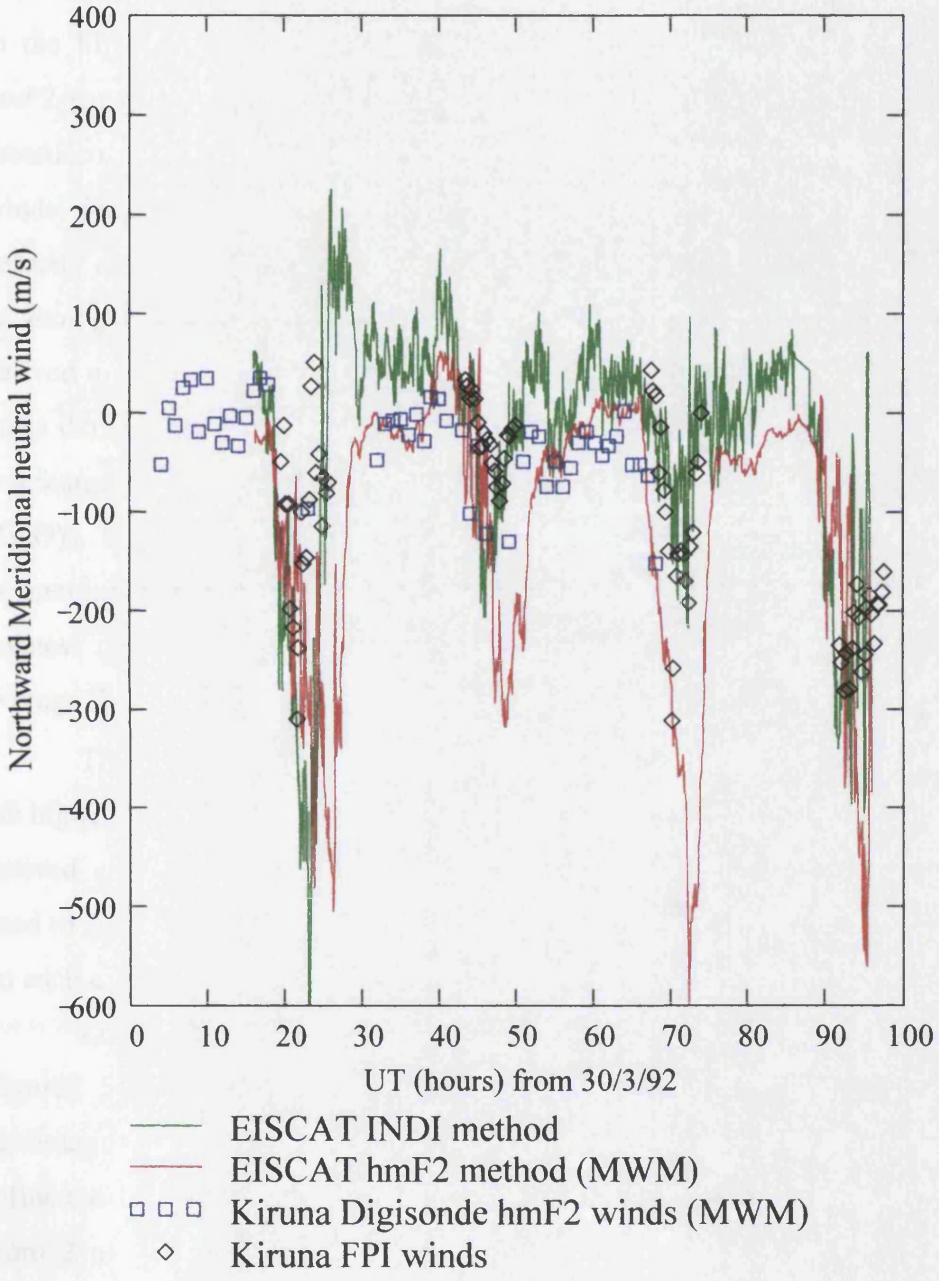


Figure 5.3 Data from four consecutive days beginning with 0000 UT on 30/3/92 showing thermospheric meridional neutral winds from four techniques: Kiruna Digisonde *hmF2* values used with MWM, EISCAT *hmF2* values used with MWM, EISCAT data used with the INDI analysis and Kiruna FPI winds. The consecutive average A_p values for these days were 12,14,13 and 6. These data contribute to the climatology for spring season at solar maximum.

There are many consistent features of the data in Figures 5.2 and 5.3. The intercomparison of winds derived from different sources of *hmF2* shows very good daytime correlation, which is encouraging. As pointed out earlier this is probably due to the higher daytime electron densities allowing more accurate determination of *hmF2* from the EISCAT measurements and the scaled parameters for the Digisonde measurements. Buonsanto et al. (1997b) have pointed out that differences between the winds derived on individual days from the same ISR experiments using the ion velocity and *hmF2* measurements could be associated with the assumption of uniform or zero gradients in the ion velocity field. In both Figures 5.2 and 5.3 the winds derived using the MWM with input from measured *hmF2* values appear to lag behind those derived using the INDI technique when reversals in wind direction occur. This is a feature of the techniques that has been noted in a previous study (Buonsanto et al. (1989)). There are also large differences in the maximum midnight southward winds as seen in Figure 5.3 with the winds derived from *hmF2* measurements producing the greatest magnitudes and largest differences to other techniques on days when the average A_p was large.

The apparent high degree of scatter seen in the INDI derived winds is due to the higher time resolution of the data used by comparison with that used for the *hmF2* derived values rather than an indication of lower accuracy. The NCAR format data used to derive the EISCAT *hmF2* values can have different time resolution depending on each experiment.

Looking at the ISR techniques in greater detail over a shorter period we see in Figures 5.4 and 5.5 examples of the influence of what has been termed the "geomagnetic history" by Aruliah et al. (1999) on the derived neutral winds, i.e. the influence of prior geomagnetic activity conditions. These figures plot the FPI winds from 2 nights with different geomagnetic histories for comparison with the winds derived from contemporaneous EISCAT CP-1 experiments using both the INDI and MWM techniques.

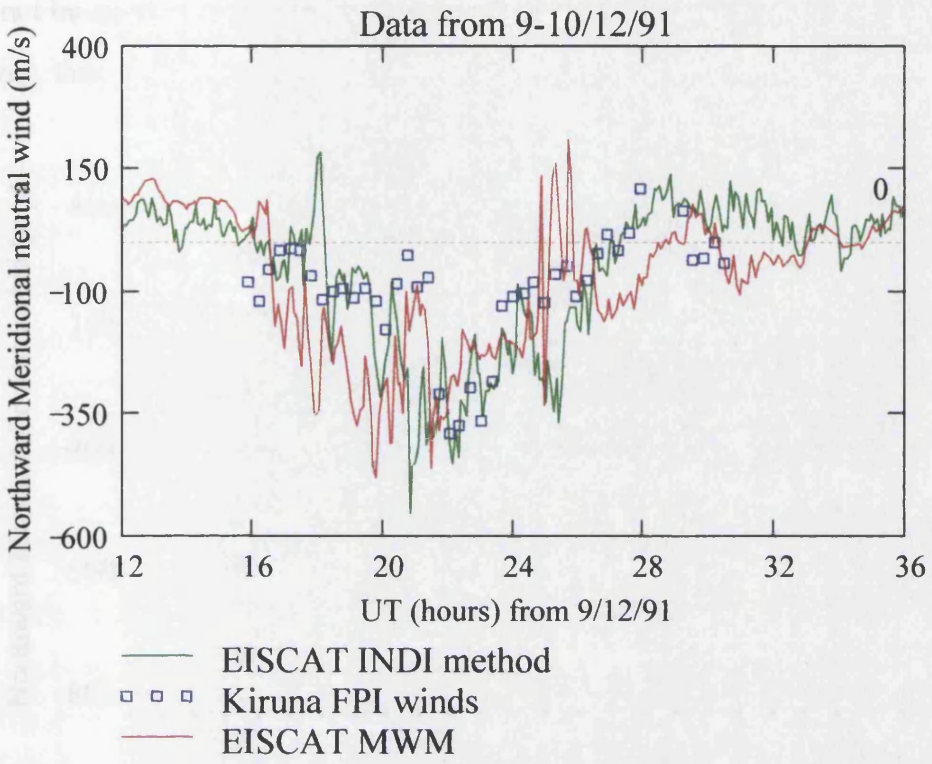


Figure 5.4 Neutral winds from 9-10/12/91, illustrating the effect of geomagnetic history – in this case a previously quiet situation when the geomagnetic activity for the previous 12 hours prior to the FPI comparison period illustrated had an average $A_p = 7$.

In Figure 5.4 the average 3 hourly A_p value during the period when the FPI winds are plotted was 16. For the period in Figure 5.5 when the FPI winds are plotted the 3 hourly A_p values average are very similar at 18. However investigating the average 3-hourly A_p values for the 12 hours prior to these periods we find that for Figure 5.4 the average was just under 7 providing previously quiet geomagnetic conditions. For Figure 5.5 the average was 12 providing previously moderately active conditions. Similarly the 24 hour average of the A_p prior to the period of the FPI comparison in Figure 5.4 were under 6 whereas for Figure 5.5 the corresponding average A_p was over 15.

With previously quiet conditions in Figure 5.4 there is reasonable agreement between the 2 EISCAT derivations of neutral winds and the direct FPI measurements. However with previously moderately active conditions as in Figure 5.5 the derived

winds are more than twice as large as the FPI winds. In other words the latter night would not be suitable for inclusion in the INDI experiment despite current A_p values being low, thus indicating the need to consider geomagnetic history.

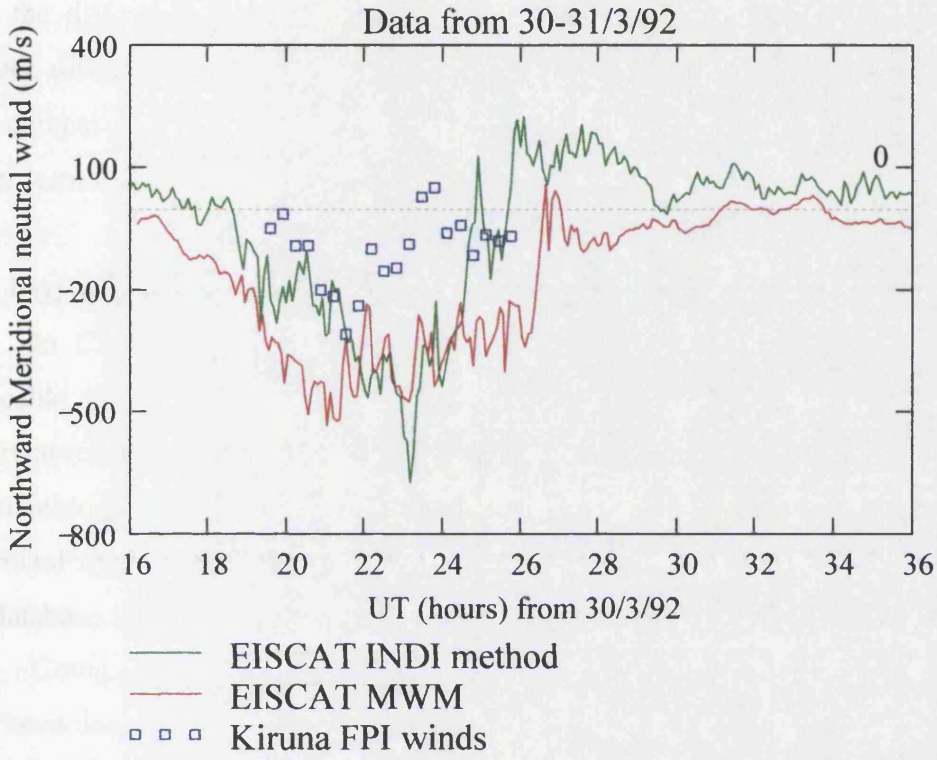


Figure 5.5 Neutral winds from 30-31/3/92, illustrating the effect of geomagnetic history – in this case a previously active situation when the geomagnetic activity for the previous 12 hours prior to the FPI comparison period illustrated had an average $A_p = 12$.

In general the geomagnetic activity will affect the ionosphere promptly and consequently the apparent winds derived from the ionospheric parameters. In contrast the direct winds measured by the FPI show the much higher inertia associated with the thermosphere. This also has an effect whereby after a period of high activity there may be a delay before the FPI winds will actually settle back down to what might be considered steady state behaviour. The ionosphere will have to a much greater degree "forgotten" about the previous activity although there may still be some influence due to the coupling with the thermosphere.

Another point to note is that there also appears to be a consistent offset of INDI winds from MWM winds. This exists for both the quiet and active cases shown in Figures 5.4 and 5.5 and indicates a difference in the way these techniques work, as

discussed earlier in this section. The ion velocity measurements used by the INDI technique produce neutral winds that are associated with a single altitude. The *hmF2* measurements used to derive the neutral winds from the MWM technique will be influenced by the full range of neutral winds in the F region. While viscosity should limit the differences, any non-uniformities in the winds at different altitudes will produce differences between the techniques. These influences may be seen in short term comparisons such as in Figures 5.2 and 5.3, but should be averaged out in the longer term climatological studies.

5.2.1 Derived climatologies

In Chapter 4 a series of climatologies developed prior to this study and applicable to the Tromso EISCAT site were presented from the WLLP, HWM empirical model winds and CTIM theoretical model. Here we present another series of climatologies derived for use in this study from three different sources. Firstly the individual nights from the EISCAT *hmF2* values used with MWM, then the Kiruna FPI database followed by the IRI *hmF2* values used with MWM.

Comparisons are more relevant for individual nights when all of the values have been derived from measurements actually made on those days, therefore none of the modelled winds (HWM, CTIM or IRI *hmF2* with MWM (IRI-MWM)) were shown on the previous plots. This allows the comparisons to be considered valid even given the possible influence of what might be termed the day to day weather of the thermosphere-ionosphere coupled system, with possible influences of gravity waves and tides as well as those influences identified in the previous section. These differences should not be manifest however in the longer term climatological comparisons, although the possibility of an average effect due to these influences emphasises the need for large databases in deriving the climatologies. Even the models based on empirical datasets can be unreliable in this respect due to the lack of relevant input data from a specific site and the smoothing effect of global averaging. The implications of this effect will be dealt with in depth in Chapter 6.

There has been a tendency in the published literature on thermospheric neutral winds to believe that any "measured" set of results are real, i.e. either direct measurements of the winds from ground or satellite borne instruments or data interpreted from ionospheric measurements from Digisonde or ISR, when comparing with model output from both empirical and first principles models. When more than

one dataset has been produced based on measurements it is often assumed that FPI results can be regarded as "true", which would seem appropriate as they are the most direct measurement, but if they are not available then one or other dataset can be favoured and discrepancies seen as a failure on behalf of the other technique (e.g. Buonsanto et al. (1997b)). The introduction of direct FPI measurements, an alternative EISCAT based technique and an alternative empirical model technique in this chapter will allow a more comprehensive comparison to be undertaken for the climatological behaviour of the meridional thermospheric neutral wind than has been previously performed for any single site. This should allow a more unbiased comparison of the individual techniques and unambiguous identification of the strengths and weaknesses of the different sources.

5.2.2 EISCAT *hmF2* with MWM climatologies

The technique of using *hmF2* values in combination with the servo theory to derive neutral winds has obvious advantages. It is applicable to ionosondes which are the most prevalent of all the instruments used for ionospheric sensing and these instruments have been used for many years so that long-term databases exist. A number of examples of this type of analysis being undertaken with ionosonde data have been published.

Buonsanto (1990) calculated thermospheric neutral winds in the magnetic meridian at F2 peak heights over a full solar cycle at mid-latitudes for Wallops Island (37.8^{o}N, 284.5^{o}E) using ionosonde data and a technique based on the servo model. Results of the study showed a seasonal variation: diurnal mean winds were southward in summer and northward in winter. The transition in the diurnal mean winds at the equinoxes was gradual, which contrasted with the abrupt transition which appeared in the TGCM results of the time. This result was important, because the solar cycle variation in the diurnal mean neutral winds was not well understood, either experimentally (because of a paucity of data) or theoretically (because of the competing effects of increased EUV forcing and ion drag with increasing solar activity). The diurnal amplitude of the winds reported was considerably larger at solar minimum than at solar maximum. A larger diurnal amplitude at solar minimum has also been reported by Duboin et al. (1988) from incoherent scatter data at St. Santin,

and also by Miller et al. (1989) who used January ionosonde data from a number of stations.

A further study (Buonsanto, 1991) calculated winds for a full solar cycle (July, 1975 – June, 1986) at a second mid-latitude station at Boulder (40.0°N, 254.7°E). In contrast to the first study the calculations excluded geomagnetically active conditions. This permitted a Fourier decomposition of the mean daily variation of winds as a function of local time to be carried out for each month without the results being affected by variable storm effects. The results were interpreted in terms of atmospheric tides. In the earlier 1990 study the effects of vertical ionization drifts due to electric fields on the servo model were ignored. This was deemed justifiable because the emphasis was on the diurnal mean winds and the diurnal amplitude, whereas the regular variations in ion drifts associated with dynamo electric fields are basically semidiurnal. In the second study, however, since the semidiurnal and higher order components were calculated, it was necessary to take account of the dynamo electric fields. This was performed by including the electric field model of Richmond et al. (1980) in the servo model wind calculations.

These examples of climatologies made up of winds derived from *hmF2* values have been presented either with no electric field correction (Buonsanto, 1990) or using modelled values of electric field (Buonsanto, 1991). Subsequently Millstone Hill ISR data have been analysed to derive neutral meridional winds from 34 experiments spanning August 1988 to August 1995 and corrected for electric field using ion velocity measurements from the same experiments (Buonsanto, 1997b). In this study however the data were not restricted to geomagnetically quiet times and the data were not presented in a climatological sense, but combined with winds results derived from ion velocities to produce a correction factor (cfac) to the *hmF2* derived winds. This was based on the findings of Titheridge (1993, 1995a) which indicated that the meridional neutral winds derived from application of servo theory were inaccurate during the sunrise and morning period because of a shift in the unperturbed F region electron density peak downward from the calculated balance height.

The results are presented in Figure 5.6 where it is seen that the largest correction was made at nighttime, with the correction producing less equatorward winds. They find, as shown in the plot, that the corrected *hmF2* winds should have smaller midnight equatorward amplitudes. For the majority (23 of the 34) of

experiments used in this study the conditions correspond to high solar activity as set out in the WLLP study. In the study the winds obtained from the ISR line of sight ion velocity measurements were regarded as "real" and the servo winds adjusted to agree with them. However the authors acknowledge that the Millstone Hill radar suffered from errors in the form of a systematic offset of around 20 ms^{-1} which varied from experiment to experiment in the winds derived from the line of sight ion velocity measurements during the period contributing to their dataset.

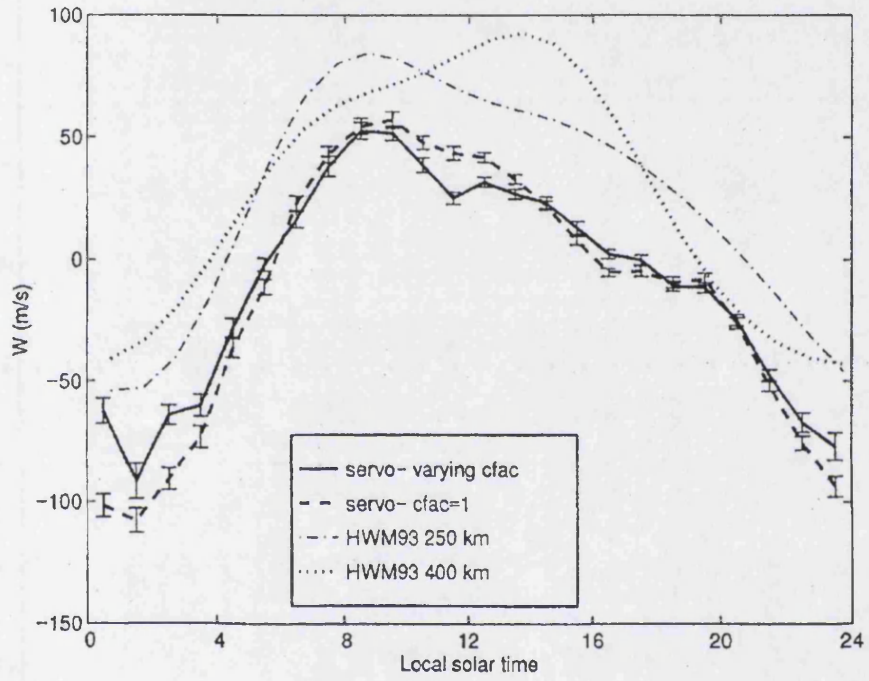


Figure 5.6 Meridional neutral winds (in the magnetic meridian positive northward) at Millstone Hill. Solid line: the mean servo model winds obtained by varying cfac to give agreement between the ISR and servo model winds for all 34 experiments. Dashed line: the mean servo model winds obtained using a constant cfac = 1. Dashed-dotted line: winds from the horizontal wind model (HWM93) for March equinox for the mean levels of solar and magnetic activity during the 34 experiments at a height of 250 km. Dotted line: HWM-93 winds obtained in the same way for a height of 400 km. (after Buonsanto et al., 1997b)

Moreover they include periods of moderate to high geomagnetic activity as measured by A_p in their dataset (14 of the 34 experiments would fail the A_p criteria set out in Buonsanto (1991)). Thus the resulting correction factor needs independent corroboration to establish whether or not the overall profile is repeated in separate datasets. The inclusion of the FLIP model in the derivation of the results using *hmF2*

with the MWM is believed to address the issues identified in servo theory calculations by Titheridge (1993,1995a) according to Miller et al. (1997).

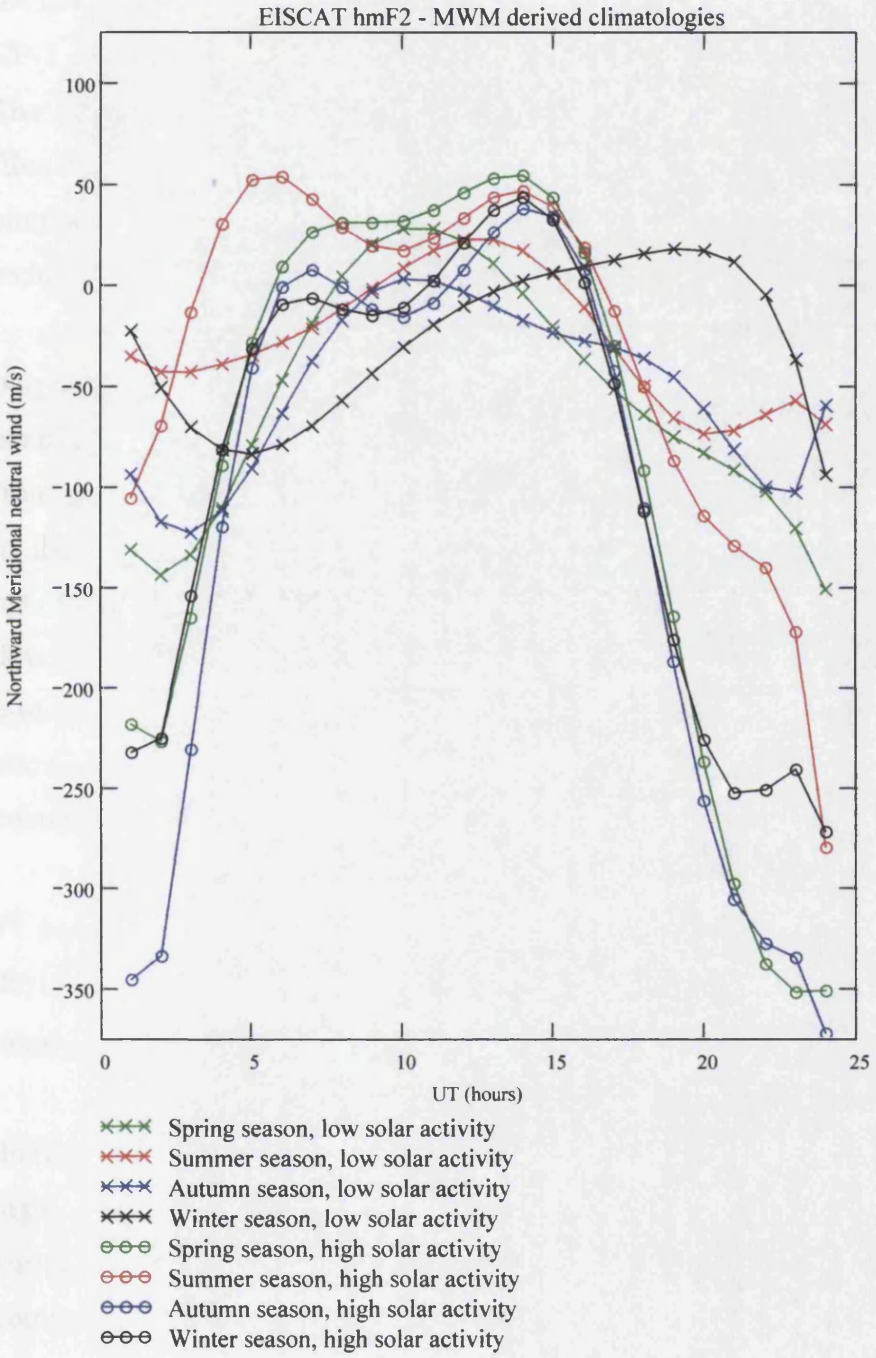


Figure 5.7 EISCAT *hmF2* with MWM climatologies

While previous climatologies have been presented of neutral winds from ISR data using techniques similar to Salah and Holt (1974), e.g. Hagan (1993) and WLLP,

here we present the first climatology of meridional thermospheric neutral winds based on ISR measurements of *hmF2* where we have used contemporaneous measurements of the electric field to correct the derived winds. The data used in this climatology are the derived neutral winds from measured *hmF2* values determined using the EISCAT CP-1 experiments with the electric field correction applied as described in Chapter 3. The data have also been restricted to exclude disturbed conditions by matching the filtering as applied in the WLLP climatology. This then provides an equivalent climatology based largely on the same EISCAT ISR database but using a different technique.

For each day or portion of a day that contributes to the climatology a polynomial fit to the data is performed and this fit is then sampled to provide a contribution to each of the hourly bins to ensure an equally weighted contribution from each night of data. This was necessary as the NCAR format result files relating to the CP-1 experiments have had varied time resolution over the course of the operation of the EISCAT radar, moving to higher resolution as the radar was developed and improved. As explained in Chapter 4 the WLLP data was made up of post-integrated CP-1 experiments, producing a fixed 3 minute time resolution, allowing them to take a simple average of the winds from each of the nights contributing to each of their bins in the different regimes.

The resultant climatologies using the *hmF2* derived winds were fitted with a 7^{th} order polynomial function and are plotted together in Figure 5.7. The standard deviations from this technique for each hourly bin in the climatology with the lowest number of contributing days range from 40 ms^{-1} to 124 ms^{-1}.

In Figure 5.7 we see that the derived set of climatologies produces a variety of diurnal variation curves. These appear relatively consistent with each other for the high solar activity situation, however for the low solar activity seasons there is more variation, with the winter season in this regime having the most idiosyncratic shape by comparison with all the other climatologies. The predominant difference between high and low solar activity plots is that at high activity there are consistently higher amplitudes in equatorward winds around midnight than are seen in the equivalent seasons for the low activity situation. There is also a more pronounced dip in the northward velocities near noon at high activity than is seen in the low activity regime. It was shown in Figure 5.6 in this section that Buonsanto et al. (1997b) found, with a

dataset from predominately high solar activity conditions, an empirical correction to the *hmF2* derived winds was required to reduce the midnight equatorward winds. This was in order to produce the best fit to contemporaneous winds derived from the ion velocity data contained in the same dataset. We see in Figure 5.7 here that the high solar activity winds display distinctly higher equatorward midnight amplitudes than those seen in Figure 4.6 in Chapter 4. This is a similar comparison of winds derived using the same ISR experiments but using different methods and displays the same trend.

5.2.3 FPI climatologies

Most climatological work on thermospheric neutral winds has been undertaken using long term data from specific FPI instruments producing direct wind measurements. Long-term FPI databases have been used to present climatologies from low latitudes (Biondi et al., 1999) using a database running from 1980 to 1990 and from the polar cap (Killeen et al., 1995) using a shorter period from 1985 to 1991. Neither of these studies has had the benefit of being able to compare with other measurements from instruments producing long term neutral winds however. Biondi et al. (1999) compare their data to TIEGCM output while Killeen et al. (1995) compare to both HWM and VSH predictions for their latitude.

The database of neutral winds from the Kiruna FPI instrument, which was described in detail in Chapter 3, has been used to produce climatologies (Aruliah et al., 1996a) and also to compare with the CTIM model (Aruliah et al., 1997). Its location also allows comparison with other data available from instrumentation in the vicinity, as demonstrated in the earlier data presented in Section 5.2 with contributions from both EISCAT and Kiruna Digisonde measurements. Previous comparisons have included the INDI experiment described in Chapter 3. The database has recently been used to investigate the evidence of an equinoctial asymmetry in the neutral winds (Aruliah et al., 1996b).

The database used by Aruliah et al. (1996b), from the first 9 years of operation of the Kiruna FPI contained 1242 nights of data and was large enough to divide into three seasons and two levels of solar activity without losing statistical reliability. The data covered the period from November 1981 to April 1990, containing observations collected between mid-September and mid-April each year. For the seasonal analysis a regime similar to that of WLLP was used with any data from within +/- 45 days of

September 21 classified as September equinox winds, within +/- 45 days of December 21 as December solstice, and within +/- 45 days of March 21 as March Equinox.

The database was also divided into two similarly sized portions in terms of solar activity. The data covered the majority of the 11 year solar cycle, so that data collected between September 1983 and April 1988 were chosen to represent solar minimum, and the data collected before and after this period represent solar maximum. The cut-off between the levels of solar activity was chosen to be an $F_{10.7}$ level of 120, which was also used by WLLP and has been used for all of the climatologies presented in this thesis.

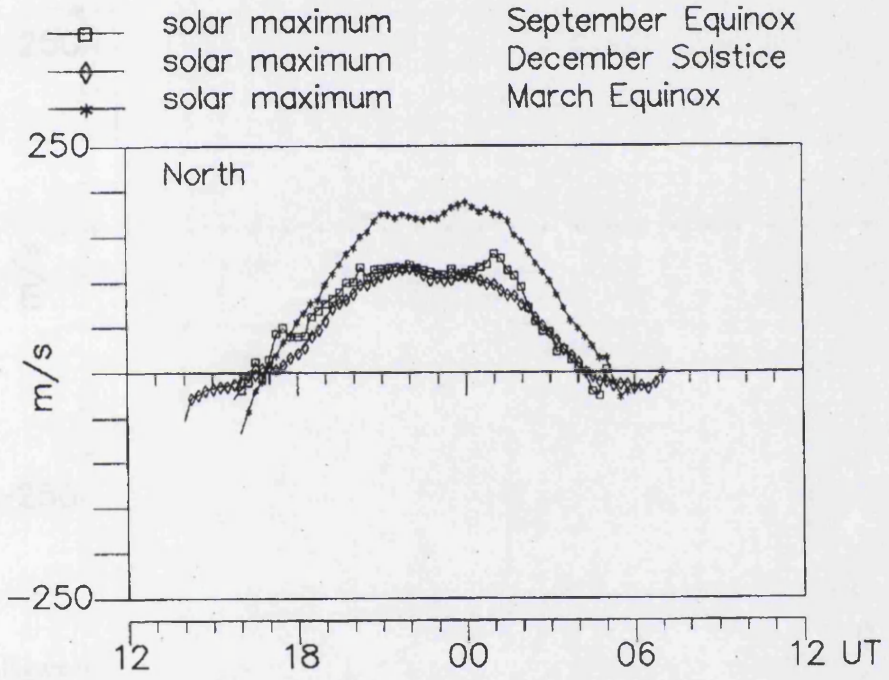


Figure 5.8 Seasonal variation in meridional thermospheric winds, positive southward, from Kiruna FPI at solar maximum (after Aruliah et al., 1996a).

Figures 5.8 and 5.9 show the evidence of the equinoctial asymmetry in the meridional winds identified in the study, at both high and low solar activity. In both figures the boxes represent the September equinox winds, the diamonds represent December solstice winds and the stars represent the March equinox winds. The size of the database also allowed the influence of geomagnetic activity to be investigated and Figures 5.10 and 5.11 show the result of restricting the data to geomagnetically quiet times ($K_p < 3$). Again the evidence of the asymmetry is apparent in both figures. This

has an important significance for the techniques which use the EISCAT data as a source of winds, which have rejected data from times of high geomagnetic activity. Aruliah et al. (1996b) attribute the equinoctial asymmetry to a solar-wind magnetospheric coupling effect that is transferred to the neutral wind through ion-neutral coupling. The presence of an equinoctial asymmetry in data where $K_p < 3$ indicates that the choice of K_p is not a sufficient selector for the quiet conditions required by the radar and ionosonde techniques.

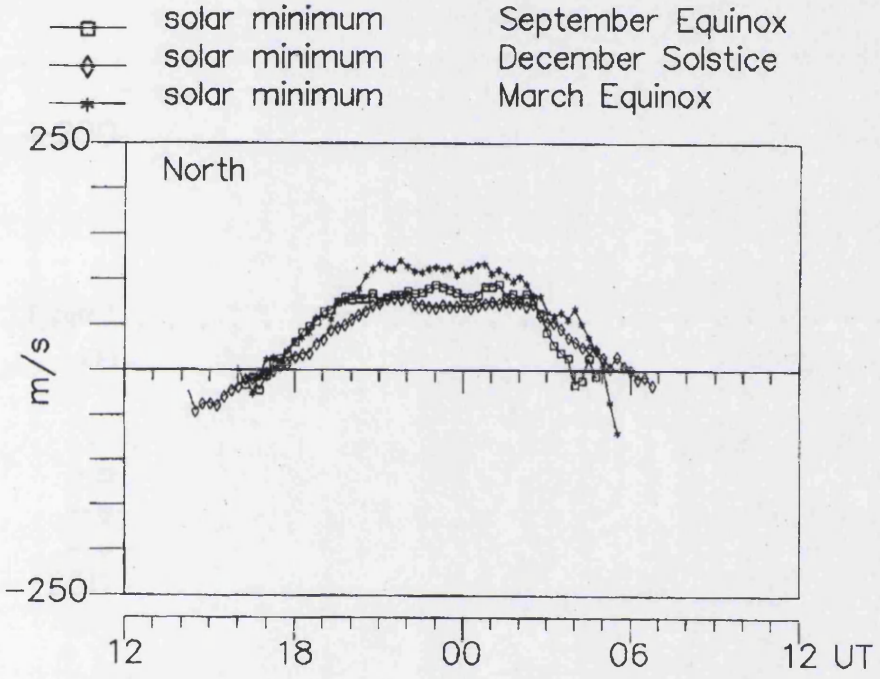


Figure 5.9 Seasonal variation in meridional thermospheric winds, positive southward, from Kiruna FPI at solar minimum (after Aruliah et al., 1996a).

In order to be able to compare the climatologies derived from the various techniques so far mentioned with direct measurements of the neutral wind a series of climatologies have been produced from the FPI wind measurement database from Kiruna which are shown in Figure 5.12. This uses all of the data available in the Aruliah et al. (1996a) study but is also extended to include the more recent data available. This leads to coverage from November 1982 to March 1998. The data have been binned according to the scheme in WLLP and provides nighttime data for comparison for three of the four seasons in both solar activity levels commensurate with the FPI data coverage limitations previously pointed out.

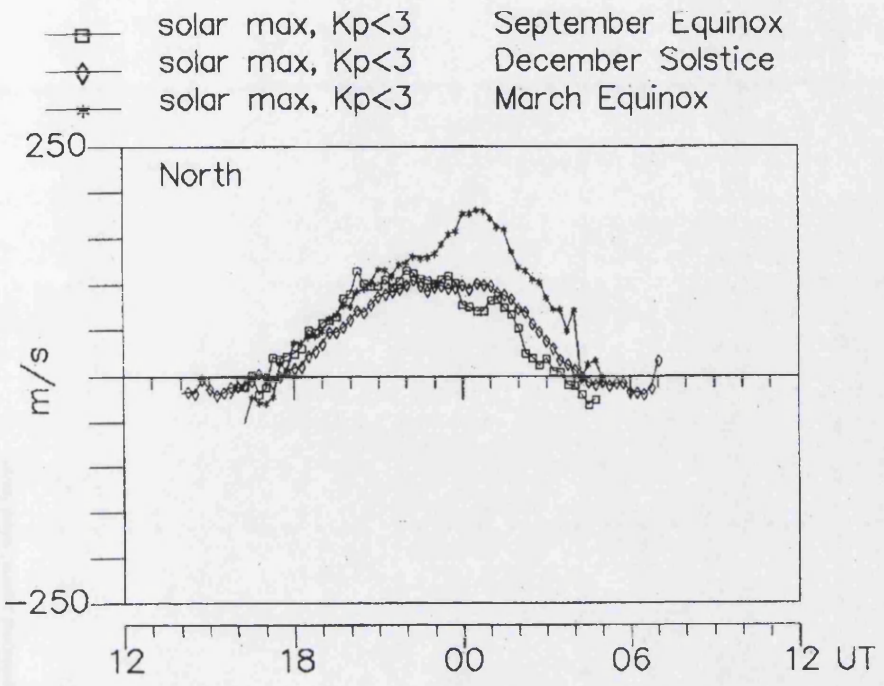


Figure 5.10 Seasonal variation in meridional thermospheric winds, positive southward, from Kiruna FPI at solar maximum for geomagnetically quiet periods only (after Aruliah et al., 1996a).

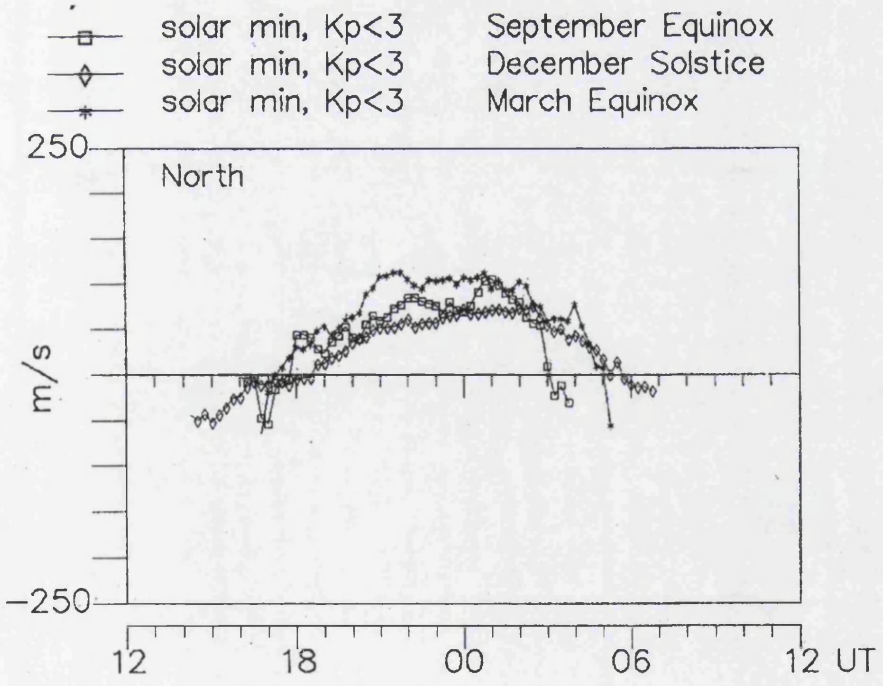


Figure 5.11 Seasonal variation in meridional thermospheric winds, positive southward, from Kiruna FPI at solar minimum for geomagnetically quiet periods only (after Aruliah et al., 1996a)

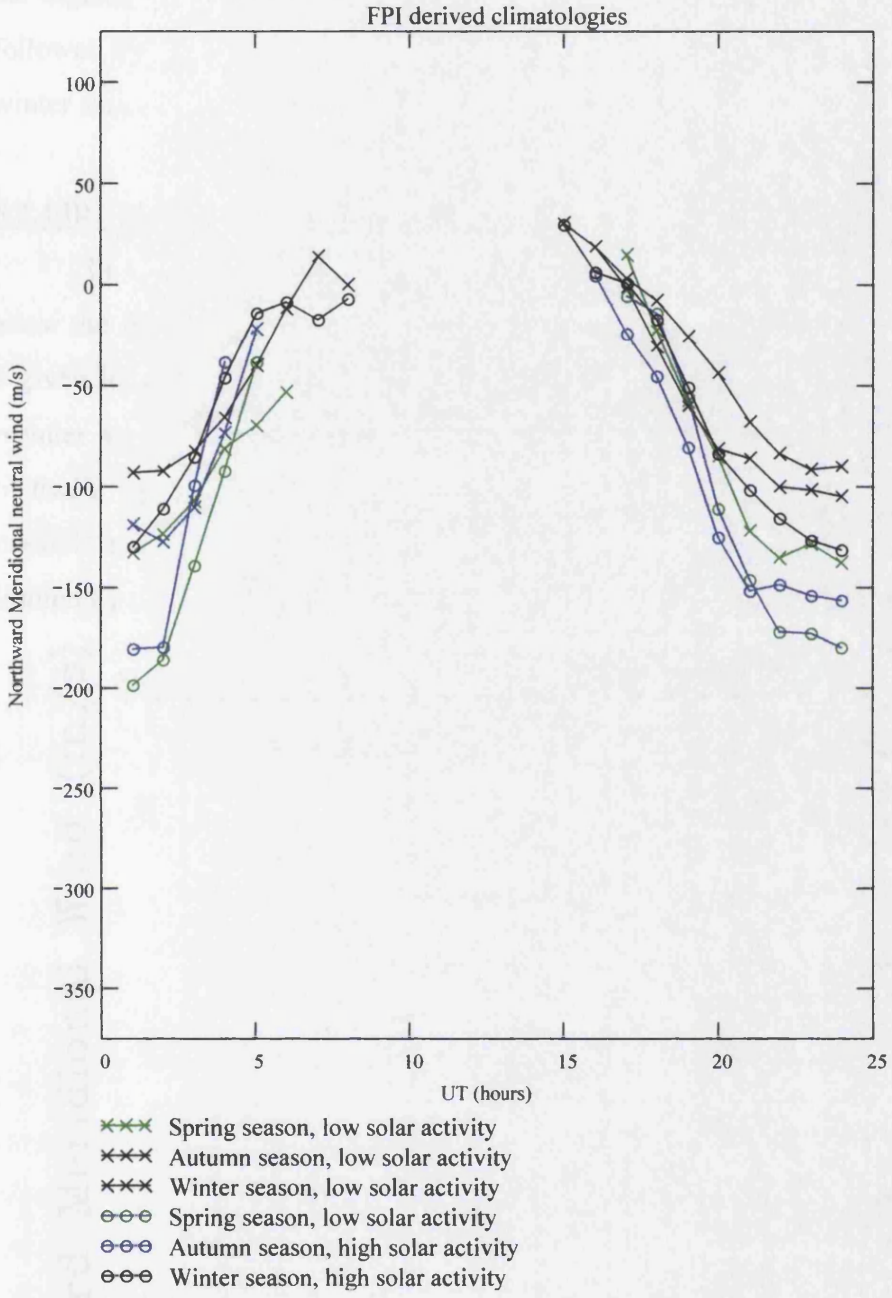


Figure 5.12 FPI climatologies

There are no summer FPI data because there is an insignificant length of dark conditions at nighttime for optical observations at this latitude in summer. There is less variability evident in the FPI winds than that seen in the EISCAT *hmF2* with MWM (EISCAT-MWM) winds over the periods covered by the FPI. The amplitudes of the equatorward winds near midnight appear to be consistently higher for the high solar activity conditions than for their low solar activity equivalents. Another point to

note is that within both of the solar activity scenarios there is a consistent pattern of the highest amplitude equatorward winds near midnight being in the spring season, followed by the autumn season and the least equatorward amplitudes being seen in the winter season.

5.2.4 IRI-MWM climatologies

In Chapter 3 it was mentioned that the MWM has been constructed so as to allow the IRI to be used as a source of *hmF2* when no measurements are available for a given location and time. This feature has been used to test results derived in this manner by comparison with both other models and against single site measurements of *hmF2* by Miller et al. (1997). Using the IRI model as a source of *hmF2* makes possible the generation of a global description of the meridional winds that is not limited by the geographic distribution of the measurements.

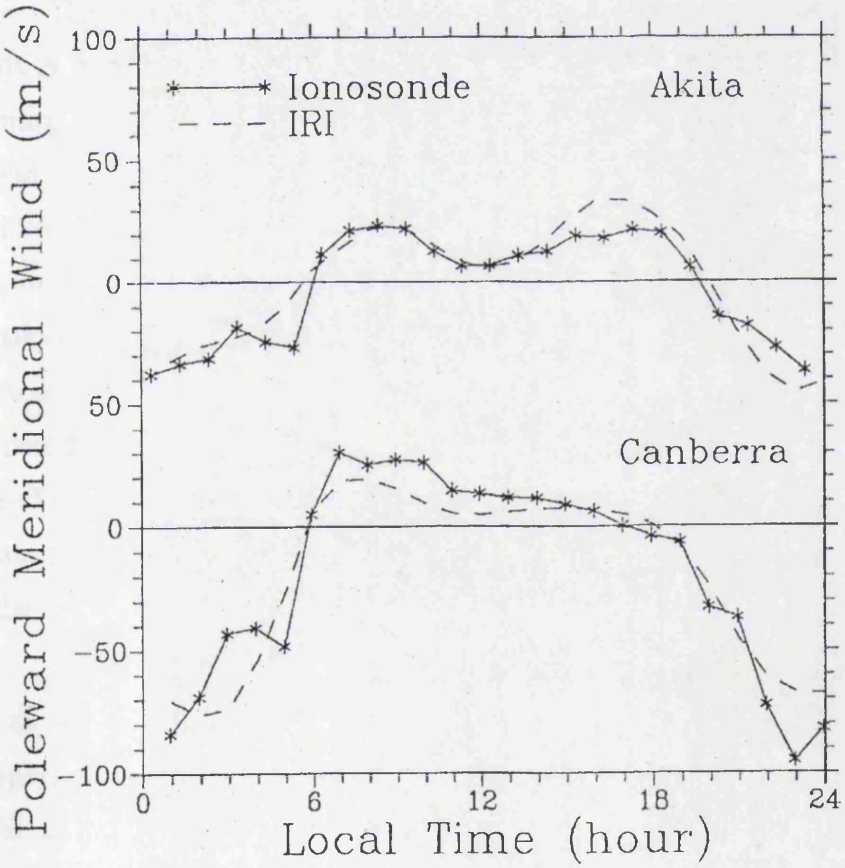


Figure 5.13 Meridional winds derived using *hmF2* from the IRI model compared with winds derived from monthly medians of ionosonde measurements. The geographic location of Akita is 40°N, 140°E, Canberra is at 35°S, 149°E. (after Miller et al., 1997).

The IRI uses a set of empirical coefficients to derive the peak height and maximum electron density of the F2 layer. The coefficients were derived from several years of ionosonde data, from a wide range of geographic locations, and their values agree closely with the median values measured each month by ionosondes. This provides a useful prediction of meridional wind values where no ionospheric measurements are available. Figure 5.13 compares winds from the IRI results with winds from ionosonde measurements from stations in Japan and Australia. The wind was plotted with poleward as positive, so that the nighttime wind is toward the equator at both stations. The solid line is the meridional wind derived from median monthly values of ionosonde parameters for March 1973. This was a moderately quiet period during sunspot minimum. The median A_p was 16 and the median $F_{10.7}$ was 99.

The poleward increase in the wind that occurs at sunrise in the ionosonde winds at both stations is caused by a minimum in the measured *hmF2*. At Akita, the morning minimum in *hmF2* is followed by a rising of the layer near noon, and a second minimum near sunset. Each of these features is reflected in the derived meridional wind speed. The noon rise in average *hmF2* is observed at many stations and, to a varying extent, at all seasons. The effect on the derived wind is that in climatological studies such as this there is a noon-time abatement in the average poleward wind. The sunrise feature may not be the result of the meridional wind only, but could be influenced by the non-equilibrium conditions in the ionosphere at that time (Titheridge, 1993, 1995a). However, Miller et al. (1997) claim that this problem should be compensated somewhat by the photochemistry that is contained in the FLIP model.

Miller et al. (1997) also compare HWM90 winds to those derived using the IRI as an input source to MWM. Figure 5.14 from Miller et al. (1997) shows a comparison of winds from the MWM with winds from the HWM90 described by Hedin et al. (1991), computed for March equinox at 1200 UT. This is a simulation of solar minimum conditions, with an A_p of 20 and $F_{10.7}$ of 100. This comparison has particular relevance as both the HWM90 and IRI are based on sets of global mathematical functions that have been fitted to data. While the satellite data set used in the development of the HWM90 is global, the ground-based data are predominantly

from the radar chain near 290° longitude. The IRI database is from ionosonde stations that are also concentrated in a few geographical regions. The result is that both models are extrapolated to regions and geomagnetic conditions that are not well represented in the data.

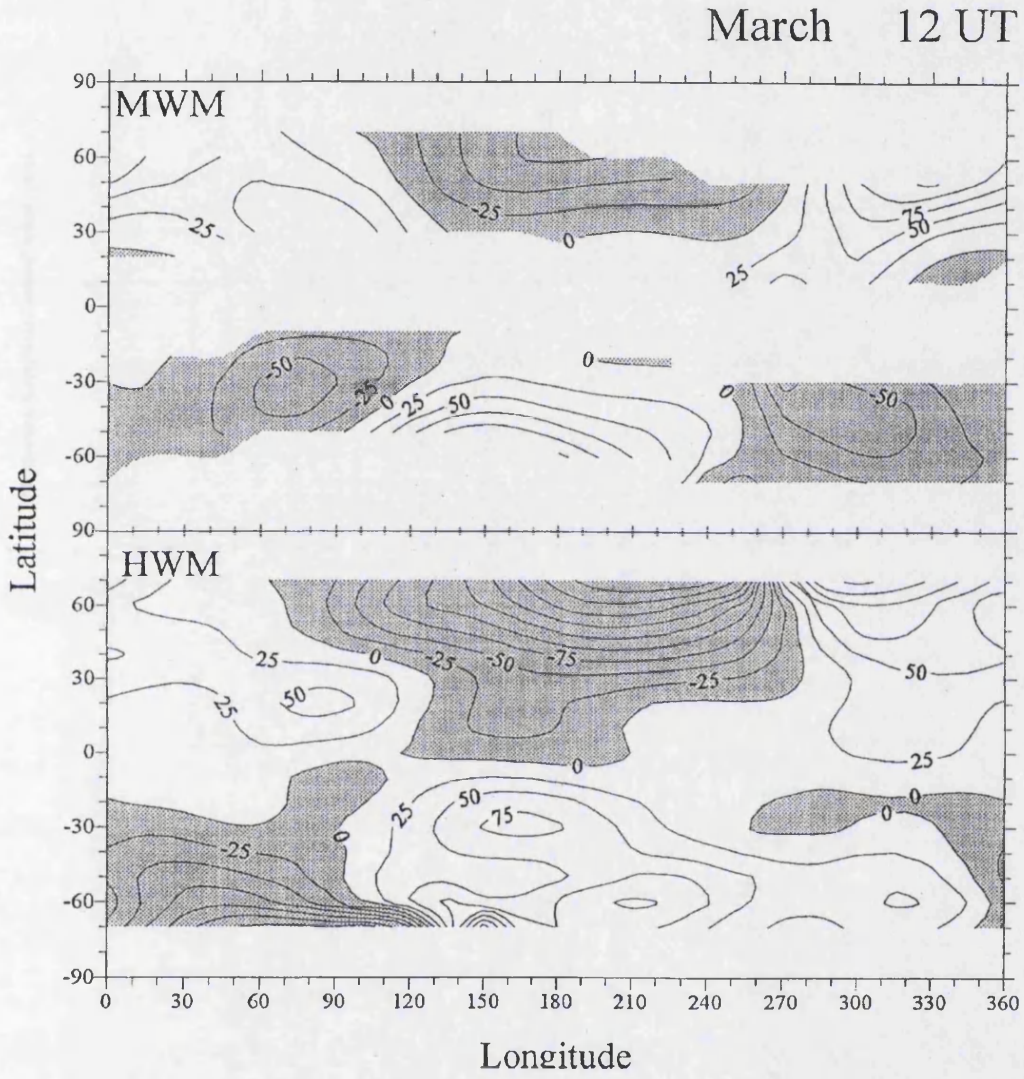


Figure 5.14 Comparison of winds from the Meridional Wind Model with results from HWM90. Regions where the FLIP model is not valid are not included in the plot. (after Miller et al., 1997).

Figure 5.14 shows that, on a global scale, the variation of the winds from the HWM90 model is similar to those of the MWM with IRI *hmF2* input. As is seen in the Figure the limitations of the FLIP model, used to calculate the "balance heights" for the MWM, restrict the MWM winds to magnetic latitudes that are more than 20° from the magnetic poles and equator.

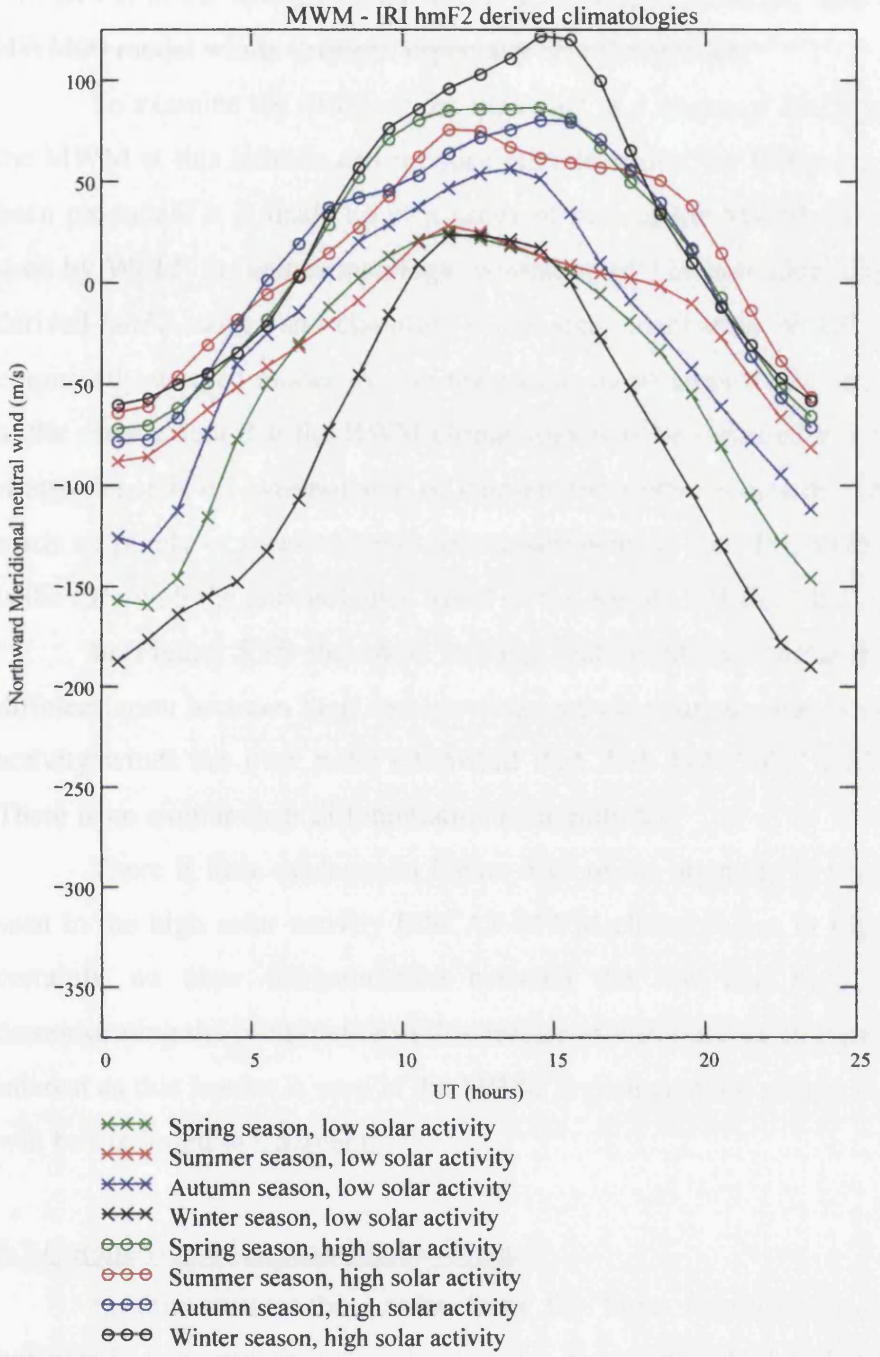


Figure 5.15 IRI-MWM climatologies

Miller et al. (1997) state that comparison at other local times and other seasons show a similar degree of consistency between the two models. Both models show the wind in general on a global scale to be more southward at the June solstice and northward at the December solstice, as expected based on the direction of the thermal gradient. However, Miller et al. point out that the difference in wind magnitudes at specific locations can be quite large. They also conclude that the major differences in

all cases is in the strength of the winds at the higher latitudes, with the high latitude HWM90 model winds stronger, especially near the solstice.

To examine the ability of the IRI to act as a source of *hmF2* measurements for the MWM at this latitude and produce reliable winds, the following climatology has been produced. It is made up of a series of runs of the MWM for each of the dates used by WLLP in their climatology, consisting of 121 individual days, using the IRI derived *hmF2* values and classified in the same manner as WLLP. As the IRI is an empirically derived model we can regard this as an empirically derived climatology, to the same extent that the HWM climatology may be regarded as empirical although it also depends on extrapolation of datasets from other locations. Therefore it is not made up purely of measurements from instruments at the EISCAT/Kiruna location as is the case with the climatologies based on the Kiruna FPI and EISCAT databases.

In Figure 5.15 the most striking feature of the climatology is the clear differentiation between high and low solar activity diurnal mean winds. No low solar activity winds are ever more northward than their high solar activity counterparts. There is no similar clear differentiation in amplitudes

There is little evidence in Figure 5.15 of the noon dip in northward velocities seen in the high solar activity EISCAT-MWM climatologies in Figure 5.7. There is certainly no clear differentiation between the low and high activity regimes demonstrating the prominence of this feature as was the case in Figure 5.7. This is of interest as this feature is seen in the MWM implementation shown in Figure 5.13 and will be discussed in Chapter 6.

5.3 Comparison of experimental techniques

In this section the results from the three measurement-based techniques presented in the previous two chapters, i.e. those from WLLP, EISCAT-MWM and FPI, are compared. These comparisons have included the standard error on the mean from each of the techniques. Figure 5.16 shows the comparisons for the low solar activity case with the top three plots comparing EISCAT-MWM to FPI climatologies and the bottom three plots comparing WLLP to FPI climatologies. Figure 5.17 presents comparisons in the same manner but in this case for high solar activity. For both Figures there are no plots from the summer season as there are no FPI data available to use in such a comparison.

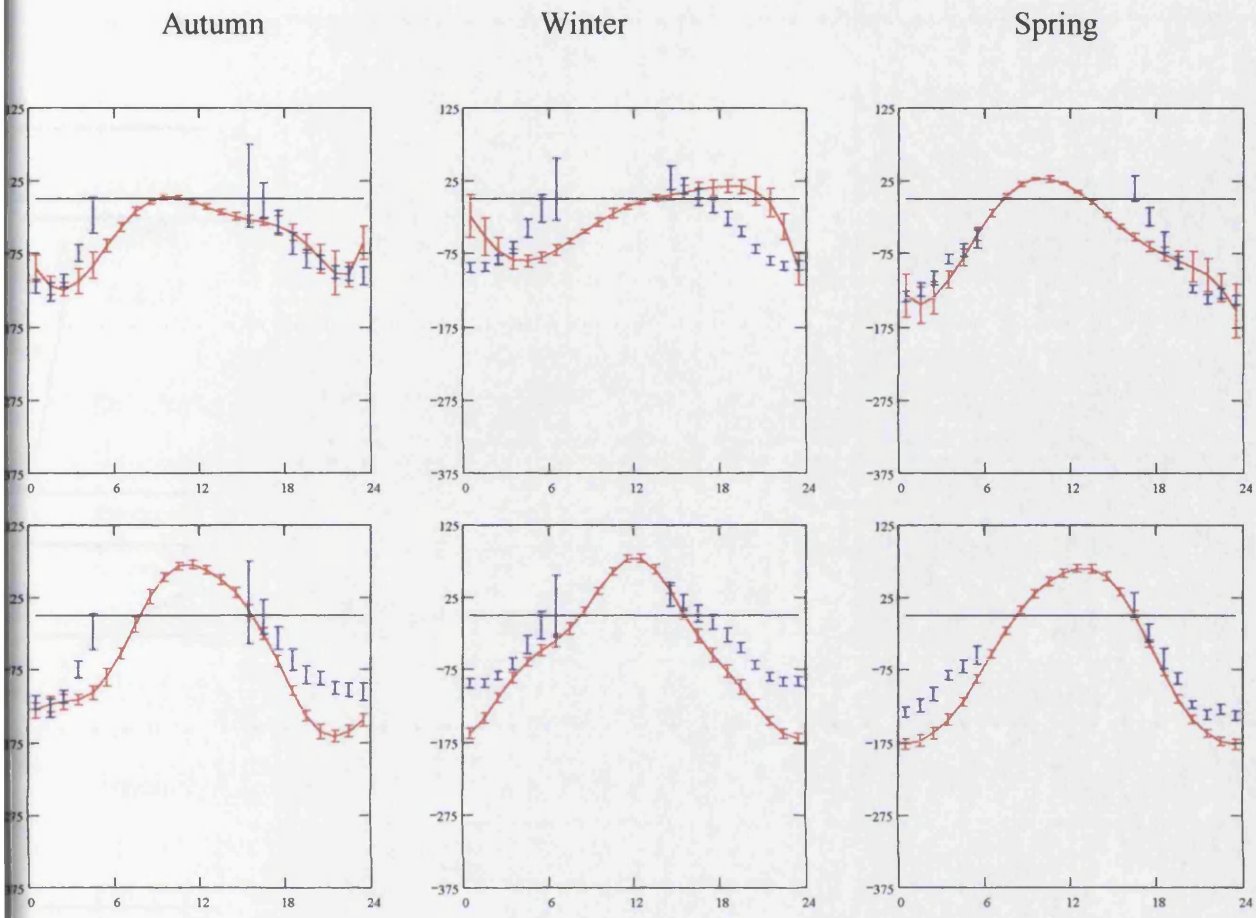


Figure 5.16 Comparison of winds for three seasons from the EISCAT-MWM technique to FPI winds (top plots) and WLLP winds compared to FPI winds (bottom plots). The y-axes in all plots denote northward meridional neutral winds in ms^{-1}. The x-axes in all plots represent UT hours. In both sets the blue points with error bars denote the FPI winds with associated standard error on the mean. The red points with lines and error bars represent the EISCAT based techniques with the standard error on the mean. In each plot the horizontal black line represents the zero wind level.

Examination of the plots reveal that the differences between the techniques are generally larger than that indicated by the standard error on the mean, pointing to significant discrepancies between the techniques. Also the standard error on the mean increases at nighttime by comparison with daytime for both of the EISCAT based techniques whereas the reverse is true for the FPI climatologies. This reflects the limitation of the FPI technique to periods of darkness and also the higher electron densities in daytime leading to greater accuracy with more good data points for the EISCAT based techniques.

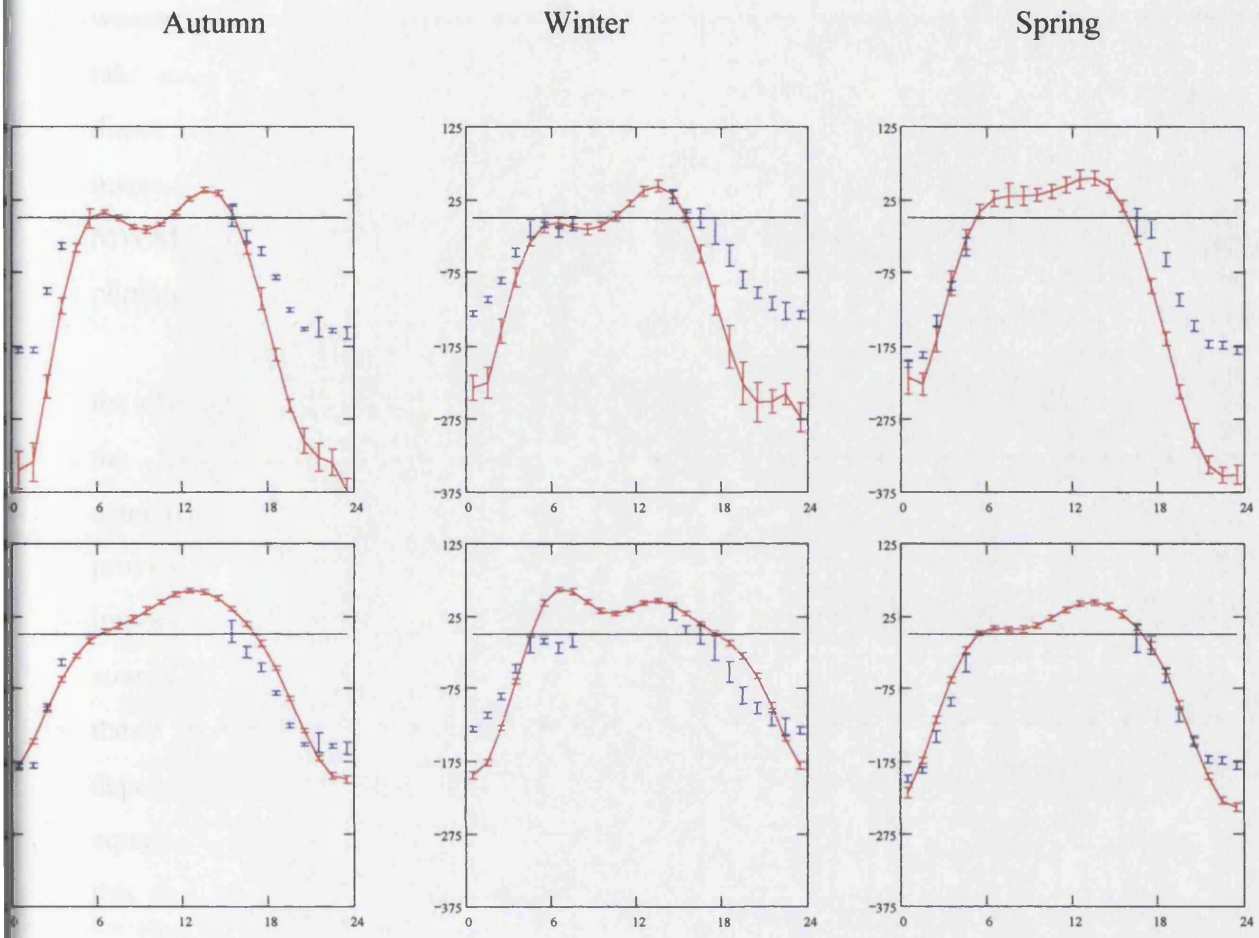


Figure 5.17 Details as for Figure 5.16 but for the high solar activity case.

In the comparisons that are presented in the next chapter the standard error on the mean will be omitted so as not to over-complicate the plots.

5.4 Summary

This chapter has presented thermospheric meridional neutral wind results from a variety of experimental techniques, initially for individual nights based on a variety of sources. These examples have been used to point out the progress made in this field and elucidate the importance of the comparisons that will be carried out in this thesis. Also presented were results from some of the largest existing single site empirical neutral wind databases to produce neutral wind climatologies. These climatologies have been developed specifically for inclusion in this thesis. In common with all of the climatologies in Chapter 4, all of the climatologies in this chapter have been plotted using the regime defined in WLLP. The background to their development and

their use in the past has been discussed. One source was the EISCAT *hmF2* values which have been used to produce a climatology of neutral winds with corrections to take account of the influence of electric fields for the first time. This was followed by direct measurements from the Kiruna FPI and the use of the database of measurements from this instrument in the past has been demonstrated. Finally the MWM used with the IRI from all of the nights that contributed to the WLLP climatology was presented.

As with the climatologies introduced in Chapter 4 the nature of the sources for the climatologies presented in this chapter have important significance. Specifically the EISCAT-MWM climatology allows a second method for neutral wind determination to be applied to the EISCAT ISR database. The IRI-MWM climatology provides a second, empirical model source of input for the MWM. Probably of most importance however is that the Kiruna FPI provides a climatology based on a direct source of winds. Some important results have been found regarding the comparison of these techniques. The clearest differences between the seasonal and solar activity dependence of diurnal amplitude, mean winds and also the amplitude of nighttime equatorward winds as derived by the three climatological techniques introduced in this chapter have been pointed out. A full comparison of all the techniques used to present neutral winds in each season and for both solar activity levels will be presented in Chapter 6.

Chapter 6 – Discussion

6.1 Introduction

In Chapters 4 and 5 a series of climatologies were presented of meridional thermospheric neutral winds for the Kiruna/Tromso site from both modelled and experimental sources. The techniques and sources in Chapter 4 had been presented in the published literature previously, while those in Chapter 5 were produced specifically for this thesis. These will be combined in this chapter to extract the important seasonal and solar activity behaviour. A discussion section will then draw together the points addressed by this thesis and put these into context. The possibilities and recommendations for further work relevant to that presented here and a summary of the work will also be presented.

6.2 Assorted combined climatologies

Each seasonal and solar activity classification is presented separately in the following sections where the results from the different climatological techniques discussed in Chapters 4 and 5 are compared and contrasted. The features of interest in each are discussed along with the important points to be noted when taking the individual classifications in combination, contrasting the seasonal and solar activity influences.

A number of different authors have presented climatological results for thermospheric neutral winds using various criteria to categorise the data (e.g. Titheridge (1991), Hagan (1993) and WLLP). The CEDAR community instigated an investigation termed Coordinated Analysis of the Thermosphere (CAT) which established criteria for the analysis of climatological averages for summer solstice, winter solstice, vernal equinox, autumnal equinox, solar minimum, solar maximum and geomagnetically undisturbed conditions (Table 6.1). Hagan (1993) used these criteria for her investigation of the mid-latitude Millstone Hill ISR derived winds but

this CAT regime has been largely ignored by later authors (e.g. Biondi et al. (1999), WLLP and Aruliah et al. (1996a)).

Table 6.1 CAT Climatological Average Analysis Criteria (after Hagan, 1993).

Geophysical Condition	Index	Range of values
Magnetically undisturbed	K_p	< 3
Solar cycle minimum	$F_{10.7}$	< 90
Solar cycle maximum	$F_{10.7}$	> 180
Winter solstice	Month	December + January
Summer solstice	Month	June + July
Autumnal equinox	Month	September + October
Vernal equinox	Month	March + April

The use of the K_p index to categorise data according to geomagnetic activity level in this manner needs to be implemented very carefully however. Hagan (1993) showed that for the winter season at solar minimum there was a significant difference in the average diurnal wind profiles at different geomagnetic activity levels. In this comparison, shown in Figure 6.1, the "magnetically undisturbed" criteria set out by the CAT, with data from three hour periods with $K_p < 3$ when the three preceding hours also had $K_p < 3$, were further constrained. This smaller dataset was made up of data from three hour periods when $K_p < 1+$ and also preceded by a period of $K_p < 1+$ for 3 hours. The comparison of the separated datasets shows a clear difference in the averaged winds. While there were insufficient data to test this effect for other seasons and solar activity conditions the fact that this effect is seen within a dataset already restricted to geomagnetically undisturbed conditions indicates the importance of this influence on the winds.

WLLP however found that their winds displayed no significant difference when split into separate categories of $A_p > 11$ and $A_p < 11$ as shown in Figure 6.2 for high solar activity. They also tested at low solar activity and came to the same conclusion. In their case however the neutral winds were derived from experiments where geomagnetically active data were excluded. This is an important factor because

local, contemporaneous data have been used to mask out disturbed conditions. Using the A_p or K_p indices to separate periods of differing geomagnetic activity is quite an insensitive method as demonstrated by the comparison of winds presented in Section 5.2 where the geomagnetic history was shown to have an important influence. The fact that WLLP still find a significant number of days included in their data with $A_p >$ 11 implies that days indicated as moderate to active by A_p may in fact display quiet time conditions. This illustrates the limitations of this method of categorisation, failing to take into account any influence of geomagnetic history.

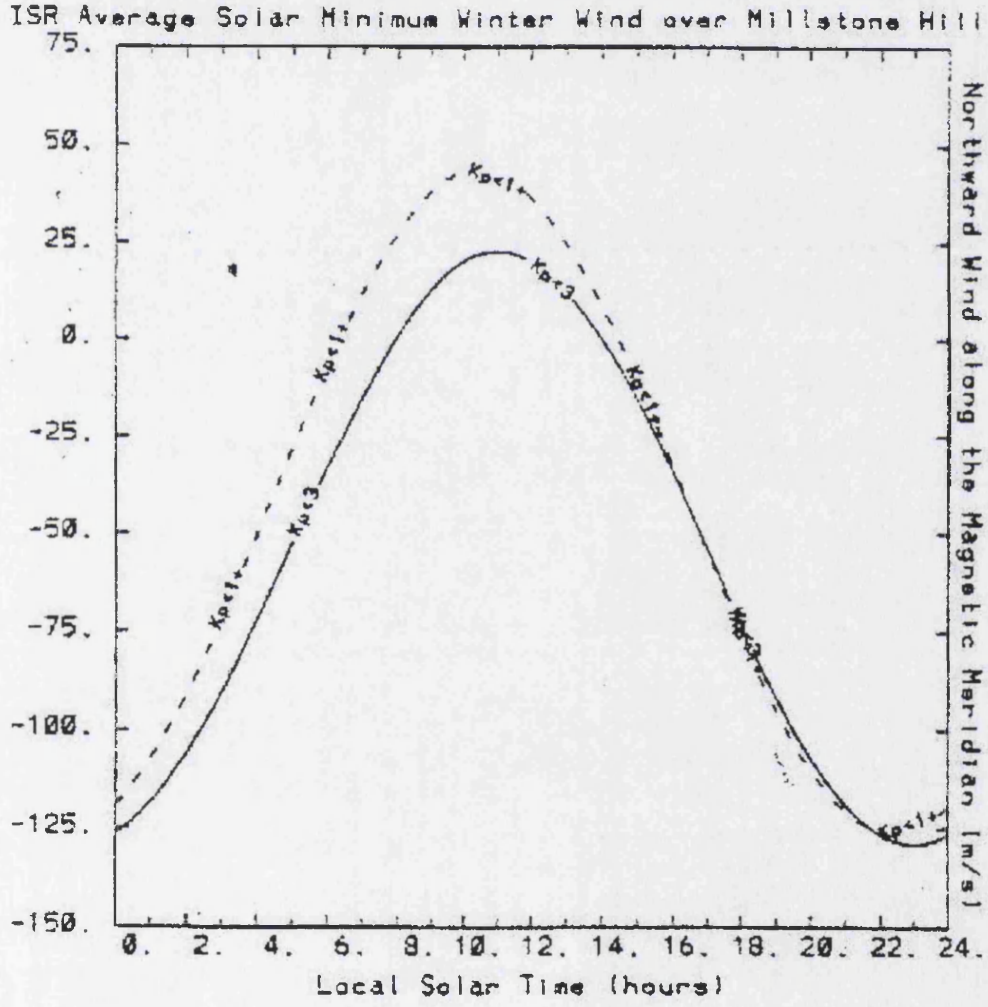


Figure 6.1 Influence of Kp level on average diurnal thermospheric neutral winds from winter at solar minimum at Millstone Hill (after Hagan, 1993).

For this thesis solar cycle and seasonal criteria were chosen to allow comparison with WLLP. While other criteria such as geomagnetic activity or IMF

conditions could be envisaged to produce further categorisation, it would impose limitations on the number of nights in each category from the experimental sources of winds and detract from the statistical significance of the comparisons. Within these comparisons the influence of short term phenomena, such as the geomagnetic history, displayed in the results from individual nights and discussed in the previous chapters, should average out and have little or no influence when examining the measurement based climatologies (i.e. FPI, EISCAT-MWM and WLLP). This is important so that we only see the average behaviour in the diurnal patterns presented.

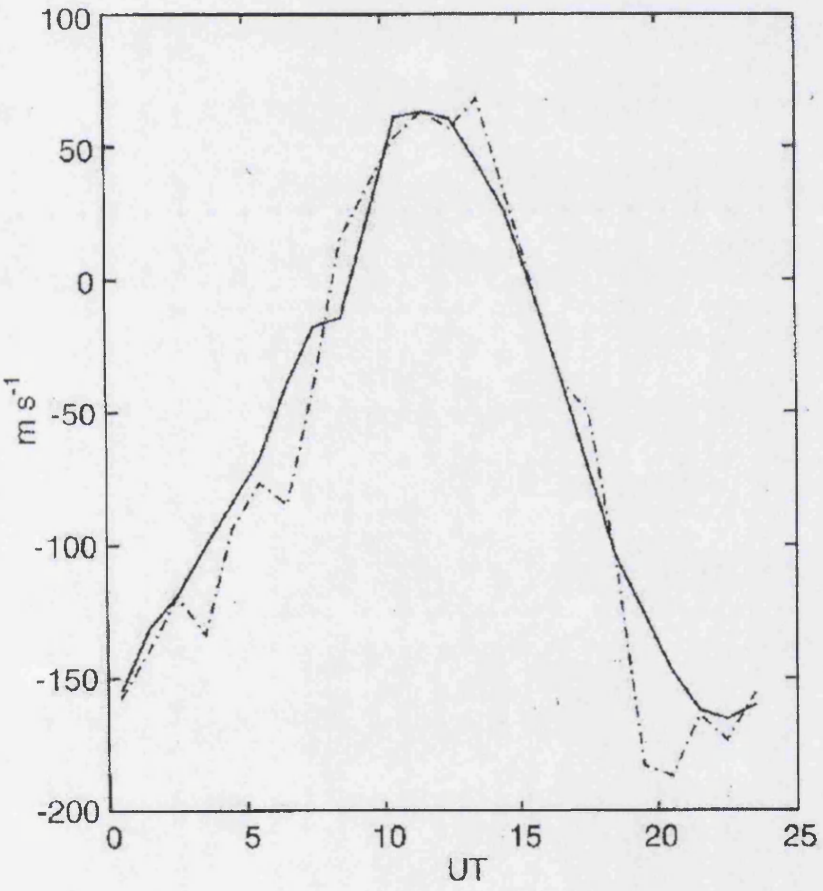


Figure 6.2 WLLP comparison of mean winds at 257 km during high solar activity. Solid line is data with $A_p < 11$, dash-dotted line data with $A_p > 11$.

The influence that solar activity may have on thermospheric winds is easy to envisage with dynamic visible features like the aurora as a reminder, and so it is a

natural candidate as a classification criteria for wind measurements. Some important seasonal effects are possible also, for instance there is a trend whereby geomagnetic conditions tend to be more active at the equinoxes as demonstrated by for example Lal (1996), which would mean there could be more influence of the auroral forcing effects at these seasons. Lal also says that there is more "geo-effectiveness" in equinox which means that for a given level of activity according to one of the indices there would be a greater influence at equinox than at solstice. Any evidence of this is important as it could call into question the use of the indices as a measure of the influence of geomagnetic activity equally for all seasons. This effect argues against the CAT criteria also and emphasises the importance of the local electric field measurements as a way of filtering data at times that are active from the ISR based techniques.

In recent years there has been much debate over the nature of the seasonal response of the meridional thermospheric neutral wind with respect to the equinoxes and whether the diurnal patterns are the same for both spring and autumn (Aruliah et al., 1996b; Biondi et al., 1999). Both Aruliah et al. (1996b) and WLLP have found that there exists an equinoctial asymmetry for the winds from the Tromso/Kiruna location. No equinoctial asymmetry was predicted by either thermospheric or ionospheric model simulations, which assume that the equinoxes are fundamentally the same, their forcing functions being symmetric about the equinoxes. Using the climatological data from separate experimental sources, and independent techniques in the case of the EISCAT data, will allow conclusions to be drawn as to the influence of seasonal effects on the datasets at this latitude.

Hagan (1993) points out that both for her study of neutral winds derived from the Millstone Hill ISR, and also from the earlier study of Duboin and Lafeuille (1992) of winds at Saint-Santin, no appreciable difference was found between the spring and autumn climatological averages. The early study of EISCAT ISR derived neutral winds by Titheridge (1991) did not consider any difference between the equinoxes, separating data by seasons into summer, winter and equinox only. This was indicative of the thinking at the time as neither the HWM87 nor the HWM90 empirical model had any difference between the equinoxes considered in their formulation. Even the more recent studies such as Biondi et al. (1999) have ignored the possibility of an equinoctial asymmetry, for their low latitude study of direct neutral wind measurements from FPI instruments at Arequipa and Arecibo.

Balan et al. (1998) find an equinoctial aysmmetry in daytime thermospheric winds from the MU radar in Japan for high solar activity conditions and the most recent mid-latitude study by Igi et al. (1999), also from Japan, has included an assessment of the differences between the equinoxes in their winds derived from ionosonde measurements from a full solar cycle. They find that the diurnal variations in each equinox are similar but that the local time of the peak winds are different at all solar activity levels and conclude that these differences have not been studied adequately to date. The mechanism proposed by Aruliah et al. (1996b) to explain the equinoctial asymmetry observed in the Kiruna FPI neutral winds is only relevant for relatively high latitudes. The work referred to above indicates that the equinoctial asymmetries may occur even at mid-latitudes and may not be restricted to the midnight southward amplitude of the winds.

6.3 Low solar activity climatologies

The results from each of the different techniques have been split into solar activity levels with the low solar activity level consisting of periods when the $F_{10.7}$ value is less than 120. When the climatologies have been restricted to low geomagnetic activity, as with the EISCAT based results here, the solar activity dependent processes should have a clearer and easier to perceive influence.

A number of effects on the winds may be expected at low solar activity. It is reasonable to think that the wind will be less northward as a result of smaller solar-driven pressure gradients. On the other hand a decrease in solar activity also means a statistical decrease in auroral heating and photo-ionization that could produce a correlation between the nighttime maximum southward wind values and the solar activity through changes in ion drag. Diurnal mean winds may also change with solar cycle, depending on the balance of the competing forces.

Quite a lot of work has been done on the solar cycle differences for the meridional neutral winds but they have mostly been at mid-latitudes (e.g. Buonsanto (1990,1991), Hagan (1993), Duboin and Lafeuille (1992) and Igi et al. (1999)). The low solar activity regime at Kiruna/Tromso may be expected to display behaviour closer to that of a mid-latitude site, given the relatively quieter geomagnetic conditions. For solar minimum at the Millstone Hill ISR Hagan (1993) suggests that auroral heating is driving southward circulation which is strong enough to compensate

for the northward circulation, driven by in-situ solar heating, most of the time, given the southward diurnal mean wind. There are many discrepancies between these previous studies however. Igi et al report that the basic behaviour with solar activity is consistent with Buonsanto (1991) as they find diurnal amplitude decreases with increasing solar activity. The details are found to be different however. They find for low solar activity that their amplitude is larger than that reported by Buonsanto (1991) and Hedin et al. (1994).

6.3.1 Spring season

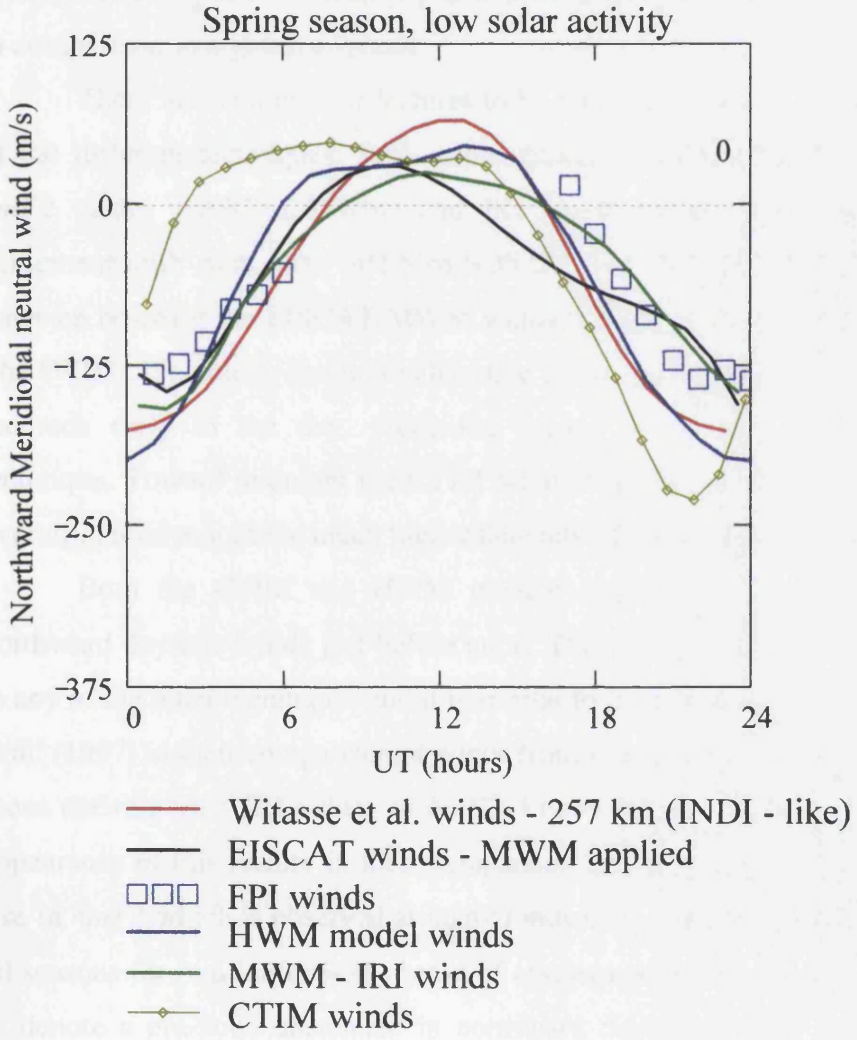


Figure 6.3 Combined contributions of spring season, low solar activity thermospheric neutral winds, positive northward, from each of the sources previously presented.

The first set of the low solar activity climatologies for comparison presented here is for the spring season. In Figure 6.3 the basic diurnal pattern is evident from all the techniques with weak daytime northward winds and stronger reverse flows at night leading to southward winds. As explained in the earlier discussion of results from a series of consecutive days in Section 5.2 these are the observations of a single site rotating with the Earth under an anti-sunward flow driven by pressure gradients caused by solar heating. There will also be an influence of auroral forcing to a greater or lesser extent depending on geomagnetic and solar activity. Since ion drag is strongest during the daytime hours when F region densities are enhanced by photoproduction, this will tend to produce wind speeds that are smaller during the day in comparison to nighttime speeds.

There are a number of features to be taken note of in terms of the comparison of the different techniques. Both implementations using MWM, i.e. with EISCAT *hmF2* values (EISCAT-MWM) and IRI *hmF2* values (IRI-MWM), are in good agreement with each other and also with the FPI data for the first half of the day. Later on however the EISCAT-MWM winds become more southward until midnight. The CTIM climatology displays values that are in marked disagreement with the other methods early in the day, producing greater northward winds than any other technique. Toward midnight the CTIM winds become the most southward giving an overall diurnal amplitude much higher than any of the other techniques.

Both the CTIM and HWM models display an abatement in the positive northward daytime winds just before noon. There is no clear evidence of this feature in any of the other techniques and it is similar to the pattern seen in winds from Miller et al. (1997) in their comparison of winds from ionosonde measurements of *hmF2* and those derived from IRI values of *hmF2*. Figure 5.13 in Chapter 5 demonstrates the appearance of this feature in their comparison and they attribute the effect to a noon rise in *hmF2* which is observed at many ionosonde stations and to a varying extent at all seasons for mid-latitudes. For ease of assessment we'll term this feature the PNA, to denote a pre-noon abatement in northward daytime winds. As noted in Section 4.3.1 WLLP noted a much smaller semidiurnal component in their winds at low solar activity. The PNA evidenced in the CTIM and HWM winds is similar to that seen in the high solar activity WLLP winds and can similarly be attributed to a higher semidiurnal component in these winds. A discussion of its significance at this high

latitude site is presented in the comparison of the different solar activity results in Section 6.6.

6.3.2 Summer season

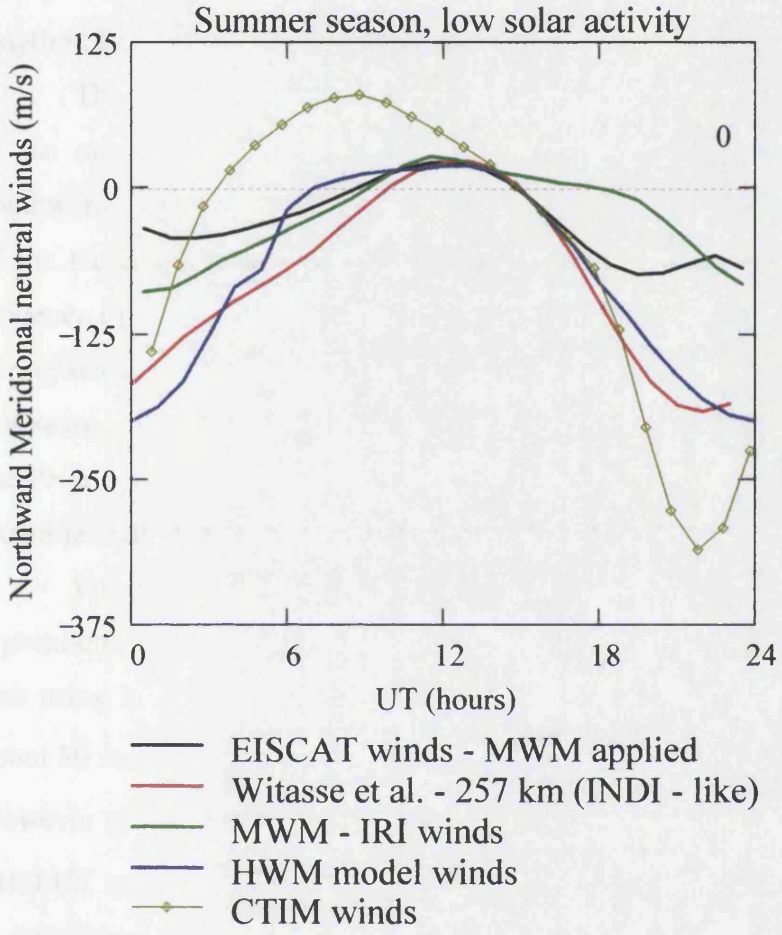


Figure 6.4 Combined contributions of summer season, low solar activity thermospheric neutral winds, positive northward, from each of the sources previously presented.

Figure 6.4 presents the low solar activity results for the summer season in which a wide variation in the nighttime wind amplitudes is evident between the techniques. There are no FPI winds for the summer season at this latitude due to the lack of appreciable periods of darkness at nighttime. Towards noon the techniques merge together displaying good agreement around the 1200 to 1500 UT period, however approaching 2400 UT there is a clear divergence with the two MWM techniques producing approximately half the equatorward wind magnitude of the HWM or WLLP winds. Of the MWM derived values the biggest differences are seen

in the second half of the day with the EISCAT-MWM winds more southward than the IRI-MWM winds as was seen in the spring case and also displaying a smaller diurnal variation by comparison with the spring season. The IRI-MWM winds display a very small overall diurnal amplitude which is significantly smaller than that seen for the spring season and produces a much lower southward maximum midnight amplitude. Daytime IRI-MWM winds are similar in the two seasons.

The CTIM model winds rapidly become too northward by comparison with all of the other techniques early in the day and these winds also become the most southward approaching midnight, again producing by far the largest diurnal amplitude of the techniques, and an increase by comparison with spring season. There is no evidence of the noon northward abatement in the CTIM winds which was seen in the spring season. The HWM winds display less pronounced daytime northward winds in this season than for spring and similar to the CTIM winds have very little evidence for the PNA compared to the spring season. As for the spring season, none of the other techniques displays any evidence of the PNA.

For the WLLP winds the amplitudes and shape are similar to spring apart from a pronounced decrease in the daytime northward wind amplitude. Titheridge (1991), also using low solar activity EISCAT CP-1 data, found an overall southward shift of about 30 ms^{-1} near noon and midnight, and 10 ms^{-1} at 0600 and 1800 UT in summer. However of his sample, about half the nights showed a large surge from 120 ms^{-1} at 2100 UT to 260 ms^{-1} at 2200 UT. While the daytime shift to more southward winds is reproduced in the WLLP winds there is no similar shift seen in the midnight wind speeds.

This could be caused by the exclusion criteria used by WLLP, as Titheridge used the mean A_p for individual days to exclude disturbed times from his dataset. WLLP actually use the electric field values derived from the CP-1 experiments as the primary criteria for rejection of disturbed times. The difference in results may indicate the shortcomings of using the mean daily A_p index as a measure of geomagnetic disturbance for the whole day and the consequent failure to take into account the geomagnetic history.

6.3.3 Autumn season

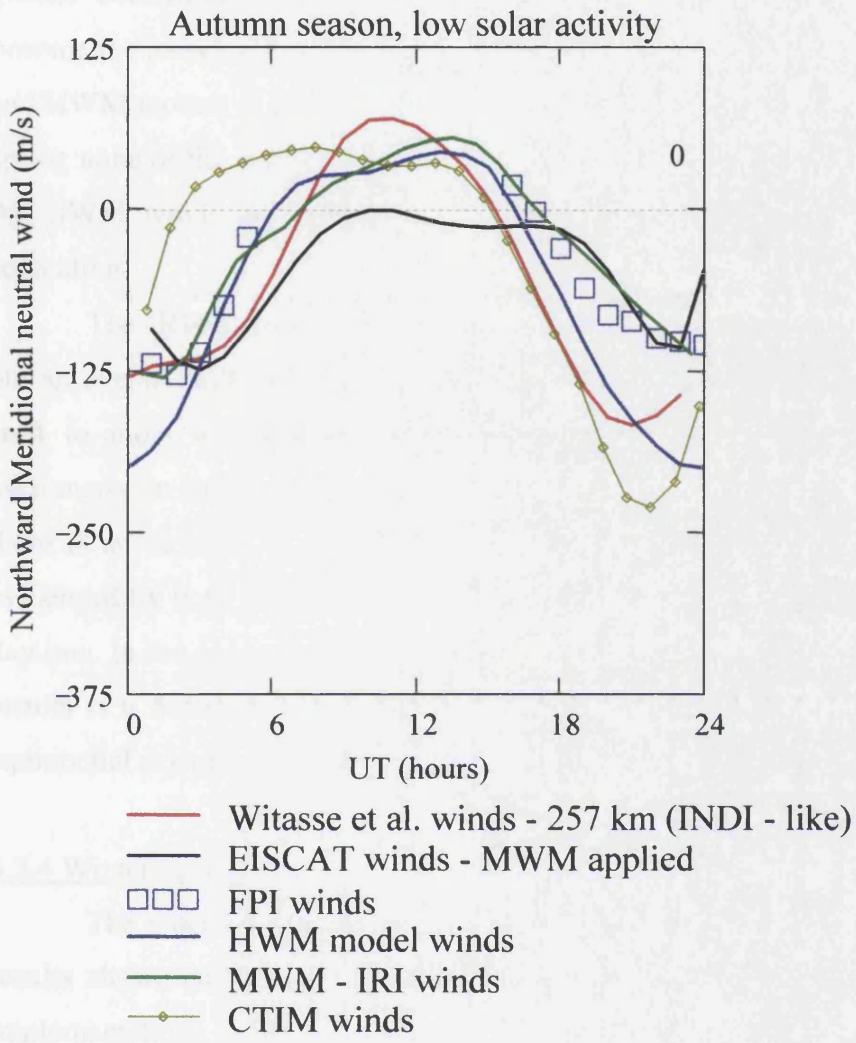


Figure 6.5 Combined contributions of autumn season, low solar activity thermospheric neutral winds, positive northward, from each of the sources previously presented.

For the autumn equinox results, shown in Figure 6.5, the comparison with both the spring equinox and summer solstice reveal interesting differences. Both implementations using MWM are in good agreement after around 1800 UT but for the earlier part of the day the EISCAT-MWM winds are much less northward, barely reaching positive northward winds and even then only for a short period. The IRI-MWM winds appear to be in better agreement with the FPI winds also.

CTIM winds in the autumn are identical to spring as there is no equinoctial difference built into the model, a feature which will also be seen in the high solar activity climatologies. Therefore the CTIM climatology again displays values that are

in marked disagreement with the other methods early in the day, producing much greater northward winds than the others, and toward midnight the CTIM winds become the most southward. Mirroring the situation in spring equinox, only the CTIM and HWM models display a pronounced PNA in the northward daytime winds. As in spring none of the other climatologies shows evidence of this feature. As with CTIM, the HWM winds are identical to spring due to the model having no equinoctial separation.

The IRI-MWM winds show a similar diurnal amplitude to the spring season but an overall shift to more northward values. The FPI winds also display an overall shift to more northward winds than in the spring season, with the equinoctial asymmetry in midnight southward winds apparent. For the EISCAT-MWM winds there is an increase in diurnal amplitude by comparison with the spring season as evidenced by both more southward winds at midnight and more northward winds in daytime. In the case of the WLLP winds the major difference compared to the spring results is a decrease in the southward midnight winds, matching the FPI observed equinoctial asymmetry, with the daytime behaviour being similar in both seasons.

6.3.4 Winter season

The most confused picture of any of the seasons so far is seen in the winter results shown in Figure 6.6 with wide discrepancies between the techniques. The implementations using MWM are not in good agreement for any part of the day, of which the EISCAT-MWM winds are significantly less southward at most times than those from IRI-MWM. The EISCAT-MWM winds appear to be in better agreement with the apparent FPI wind diurnal variation amplitude than any of the other techniques. The phase of the EISCAT-MWM variation however is very different and also differs by comparison with the phase of all the EISCAT-MWM results in the other seasons. While this season produces the least consistent set of phases when comparing the techniques, these EISCAT-MWM results are the least typical. There is a smaller diurnal amplitude than either of the equinox seasons but about the same as that of the summer season. The IRI-MWM winds display the largest diurnal amplitude in winter, compared to the other seasons, and also the greatest southward amplitude at midnight. The FPI winds in winter have a lower southward midnight amplitude than either of the equinox seasons.

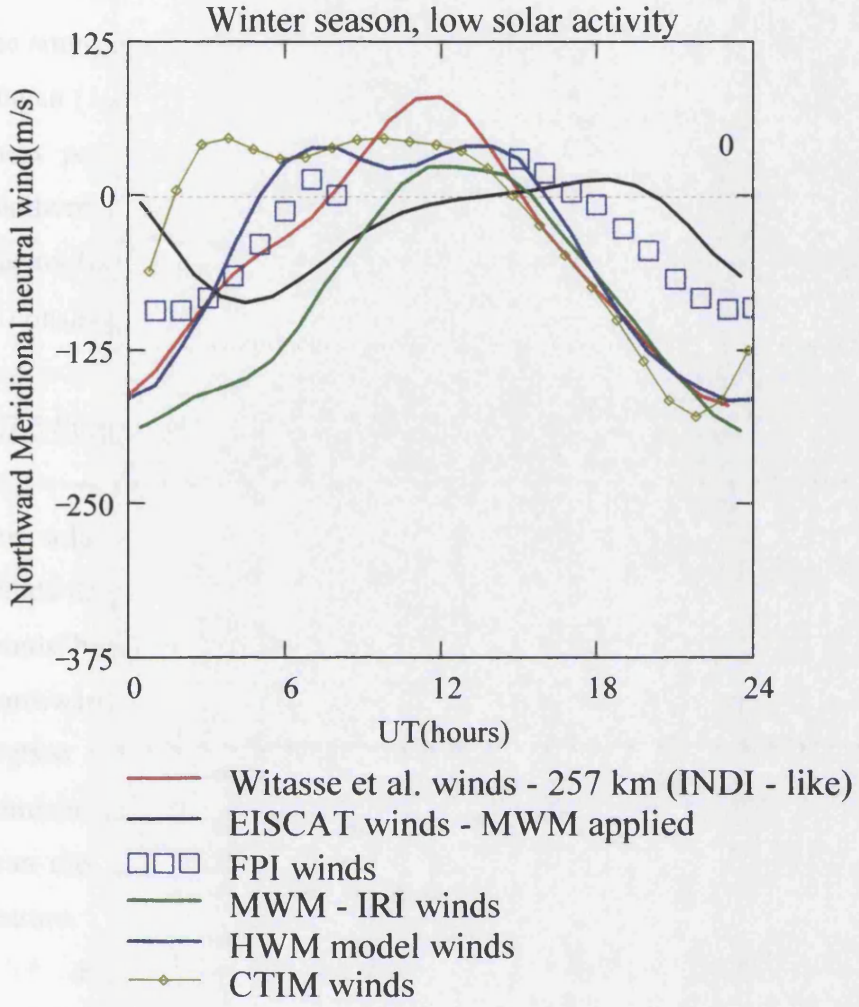


Figure 6.6 Combined contributions of winter season, low solar activity thermospheric neutral winds, positive northward, from each of the sources previously presented.

For the CTIM winds we see a much smaller diurnal amplitude than for the equinox or summer seasons. The evidence of the PNA is however greater here in the winter season than in any of the other seasons. The CTIM climatology displays values that are more northward than the other techniques early in the day but from about 1500 UT until 2300 UT the CTIM, HWM, WLLP and IRI-MWM winds are all in very close agreement. Again, only the CTIM and HWM models display a PNA from northward winds with little evidence of a similar feature in any of the other climatologies. There is no great difference in the HWM amplitude in winter by comparison with the other seasons but there is a much more pronounced PNA than in the other seasons.

The WLLP winds for winter have a similar midnight southward amplitude to the summer season but have higher amplitude northward daytime winds. As noted by Hagan (1993) for mid-latitudes, the daytime northward winds are strongest as well as most persistent during the winter because although the subsolar point is in the southern hemisphere, the northern hemisphere aurorally driven circulation is weak due to darkness. Again there is little evidence of any PNA in the WLLP winds, which is consistent with the other seasons.

6.3.5 Summary of the low solar activity comparisons

A number of clear trends and features are worth noting from the collection of low solar activity climatologies within each of the techniques. The EISCAT-MWM winds display similar diurnal amplitudes for summer and winter although the winter winds have a pronounced phase shift by comparison with the other seasons with northward winds appearing approximately 6 hours later. The spring winds have the highest diurnal amplitudes with the autumn winds having similar amplitude to both summer and winter. The spring winds also have greater southward midnight winds than the autumn winds. None of the seasons has any clear signature of the PNA feature.

For the WLLP winds there is little difference in overall diurnal amplitude through the seasons although the summer winds have the smallest daytime northward amplitude and winter winds the largest. Also there is a difference between the southward midnight maximum winds at equinox, with the autumn winds having smaller amplitude than those in spring. There is very little evidence in the WLLP winds of any PNA feature in any of the seasons.

The FPI winds display the previously reported equinoctial asymmetry with more southward winds during the night in spring compared to those in autumn. Aruliah et al. (1996b) attribute the existence of the asymmetry to asymmetric auroral forcing between the equinoxes. The winter winds display the least southward midnight amplitudes of the three seasons for which data is available.

For the IRI-MWM winds there is a trend of decreasing diurnal amplitude from winter through the equinoxes and into summer. No evidence of the PNA feature is seen in any season. There is also a noticeable difference in the equinoxes in that there is a shift to more southward winds in spring compared to those in autumn.

The HWM winds display consistent midnight southward amplitudes through all the seasons. Apart from the low daytime northward winds in summer the diurnal amplitudes are also consistent. The evidence of the PNA feature however increases from summer through the equinoxes to winter. There is no possibility of seeing any asymmetry between the equinox winds at any time as the model contains no separate consideration of the spring and autumn equinoxes.

For the CTIM winds we see that there is a consistent increase in the diurnal amplitude of the winds from the winter season through the equinoxes to the summer season which has the largest amplitude. The evidence of the PNA grows in the opposite direction to the amplitudes with the least evidence in the summer winds and the winter season having the most evident feature. This is similar to the progression in the HWM winds.

A number of clear trends emerge when examining the different climatologies together in the low solar activity regime. Clearest is that all of the techniques show that there is an equinoctial asymmetry in the midnight southward winds with autumn winds having less amplitude than the spring winds except for CTIM and HWM which is explained by the lack of consideration of equinoctial differences in the model formulations. While there aren't daylight observations to establish evidence for the PNA in the FPI winds it is clear that the EISCAT based techniques and IRI-MWM don't show any evidence of it, whereas the CTIM and HWM results do. With neither of the experimental techniques indicating the feature is present at low solar activity this must be seen as a failure of the models.

A confused picture emerges in considering the diurnal amplitudes. The techniques with clear diurnal amplitude differences, CTIM and IRI-MWM, have contradictory trends. Assuming the FPI results are a measure of the real dependence indicates that CTIM is more likely to be correctly predicting the dependence.

Low solar activity conditions at the Kiruna/Tromso site may be expected to produce the best comparison to mid-latitude conditions. In this case it is useful to compare with the large number of mid-latitude studies which have been carried out. Khachikjan et al. (1997) found that for mid-latitude calculations at 53^{o}N in summer time and at low solar activity, noon meridional winds derived from IRI-86 are directed towards the equator (southward) whereas noon zonally-averaged meridional winds predicted by the HWM90 model are always directed to the pole (northward). In

this study the noon winds from both HWM and IRI-MWM techniques are very similar but the diurnal mean wind from HWM is much more southward than that from IRI-MWM. As the implementation of the IRI-86 output is different to the IRI-MWM results presented here it is unclear as to whether this is a systematic difference due to the different implementations or whether it is a physically significant difference between the high and mid-latitude sites.

Overall Khachikjan et al. find that the IRI-86 diurnal variations vary more strongly with season than those predicted by the HWM, which is certainly in keeping with the results presented here. At low solar activity there should be less influence of auroral forcing at mid-latitudes than would be the case at high solar activity. The diurnal variation in the winds should then be larger at a higher latitude site such as Tromso/Kiruna than for the mid-latitudes. This would indicate that the HWM daily mean wind and diurnal amplitude variations are probably not large enough and that the IRI-MWM diurnal amplitudes should have greater variation also. These empirical models may be averaging out such variations or lacking sufficient data coverage to model them.

Another mid-latitude study by Igi et al. (1999) find that there is a strong seasonal difference in diurnal amplitude at low solar activity with the maxima 1 month after solstice and the minima 1 month after equinox. These mid-latitude results match those found from previous work for diurnal amplitude at other mid-latitude sites (Buonsanto 1990, 1991; Duboin and Lafeuille, 1992), the wide longitudinal coverage provided by these studies giving confidence to the findings. The opposite trend is found here in the results from the EISCAT-MWM technique. The IRI-MWM winds have maximum diurnal amplitude in winter but minimum amplitude in summer which indicates that differences using the *hmF2* techniques are not just confined to the individual techniques here but also with results from another site, albeit separated in latitude.

Hagan (1993) found that at low solar activity the largest signature of seasonal variability in the climatological average winds from the Millstone Hill ISR was seen in the comparison of the solstice results. Examining the comparable WLLP results here we see that the extremes of daytime northward amplitudes are established by the solstice winds but that the maximum nighttime southward winds and overall diurnal profiles are otherwise very similar between the seasons.

The data used from EISCAT were chosen to be from geomagnetically quiet periods, using the electric field measurements, in order to comply with the criteria for applying both the WLLP (ISR ion velocities) and MWM (servo theory) techniques, whose assumptions rely on conditions being close to an equilibrium situation. However Lal (1996) postulated a bias towards more geomagnetically active times at the equinoxes compared to the solstices. This would suggest a trend towards more southward diurnal mean winds at the equinoxes by comparison with the solstices. Of the measured winds the EISCAT-MWM and FPI winds appear to display evidence of this effect. There is little evidence of this effect in the WLLP winds at low solar activity however. As the same criteria have been used in the exclusion of disturbed conditions for both the EISCAT-MWM and WLLP results, and these techniques use the same raw data, this is a clear difference between the individual techniques. None of the modelled winds shows clear evidence of the effect.

6.4 High solar activity climatologies

The results for the high solar activity level consist of periods when the $F_{10.7}$ value is greater than 120. As with the low solar activity case, the restriction of the measured climatologies to low geomagnetic activity, as with the EISCAT based results here, should reveal a clearer influence of the solar activity dependent processes.

The increased solar activity by comparison with the previous section will lead to an increase in the auroral forcing at high latitudes but also to an increase in the solar in-situ forcing which can be thought of as competing forces. The balance between these influences at this latitude will be compared to the findings of previous mid-latitude studies to demonstrate the important differences.

Hagan (1993), for example, found that the magnitudes of the nighttime winds during solar maximum are markedly smaller than the nighttime solar minimum wind speeds. In contrast, the daytime winds at solar maximum are stronger than those during solar minimum. Also the in-situ solar forcing is dominant during solar maximum daytime conditions when there are northward winds during all seasons. During solar maximum summer when the northern hemisphere high latitude densities are enhanced by photo-production, the solar in-situ circulation is relatively weaker than it is during the other seasons. In this case, the in-situ solar forcing and the

influence of auroral forcing from higher latitudes appear to cancel one another in the early morning and late afternoon over Millstone Hill ISR and the observed winds hover near zero during these times.

In summary Hagan finds that without regard for the seasonal factors, solar in-situ forcing appears to play a more dominant role for the solar maximum winds in comparison to the solar minimum winds. This result suggests that quiet time auroral heating increases with solar activity were relatively smaller than solar in-situ heating increases between 1984 and 1990 for this mid-latitude site.

Igi et al. (1999) observe a decrease in diurnal amplitude with increasing solar activity at their mid-latitude site in Japan. They believe that the larger ion drag must play an important role in restraining the amplitude at high solar activity and they also find that within the high solar activity regime that the largest diurnal amplitudes are in winter and the minimum is in the spring equinox.

The recent study of the equinoctial asymmetry using Kiruna FPI winds by Aruliah et al. (1996b) finds that the equinoctial asymmetry difference is greater with high solar activity and that overall higher solar activity gives higher wind velocity magnitude. This runs contrary to the findings of the previously introduced mid-latitude studies and indicates that the balance of the auroral forcing and the solar heating at this latitude is quite different to that seen at the lower latitude sites. At high latitudes the influence of an expanded auroral oval, as would be expected for high solar activity conditions, is for the winds to become more southward and also for the diurnal amplitude of the wind to increase (Titheridge, 1995b).

6.4.1 Spring season

In the first of the results at high solar activity, shown in Figure 6.7, we see that most techniques display pronounced evidence for the PNA for the spring season. A prominent feature of the previously reported equinoctial asymmetry was that the effect was clearer during periods of high solar activity than for low solar activity.

The IRI-MWM derived winds are significantly more northward than all the other techniques from midday to early in the morning and in significant disagreement with the EISCAT-MWM derived winds. EISCAT-MWM winds are the most southward of the techniques from 1800 UT to 0300 UT. Both HWM and the WLLP winds appear in good agreement with the FPI winds. CTIM is the most northward

technique from around 2300 UT to 0600 UT and the winds appear to reverse from southward to northward during the night earlier than any of the other techniques.

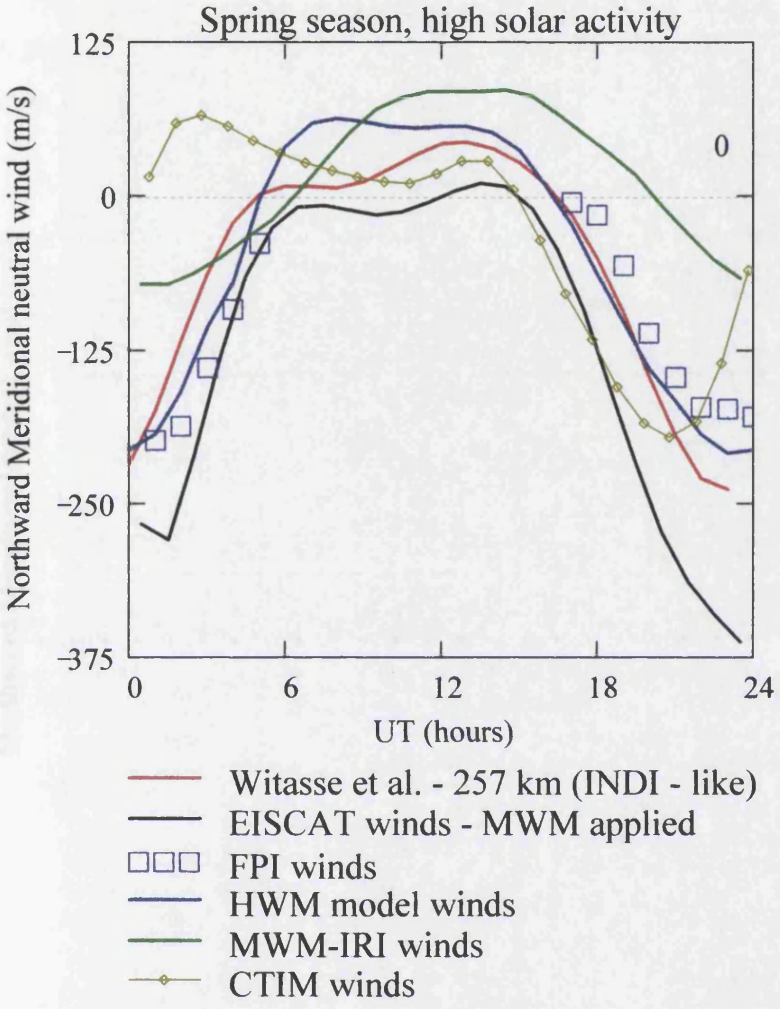


Figure 6.7 Combined contributions of spring season, high solar activity thermospheric neutral winds, positive northward, from each of the sources previously presented.

6.4.2 Summer season

Figure 6.8 shows the results for the summer season at high solar activity and we see that similar to the spring season most techniques display pronounced evidence of the PNA. Also repeating the trend seen in the spring season, the IRI-MWM derived winds are significantly more northward than all the other techniques from midday on to early in the morning and noticeably the least consistent compared to the other techniques. The EISCAT-MWM and WLLP techniques appear to agree quite well apart from a discrepancy in the EISCAT-MWM winds approaching midnight when there is a marked southward peak. For the IRI-MWM winds we see a similar diurnal

amplitude in summer compared to the spring winds and also a similar small degree of evidence for the PNA compared to spring. For the EISCAT-MWM winds we see a much smaller diurnal amplitude in summer compared to the spring winds and also more pronounced evidence of the PNA than in spring. There are no FPI winds for the summer season at this latitude as was the case for low solar activity.

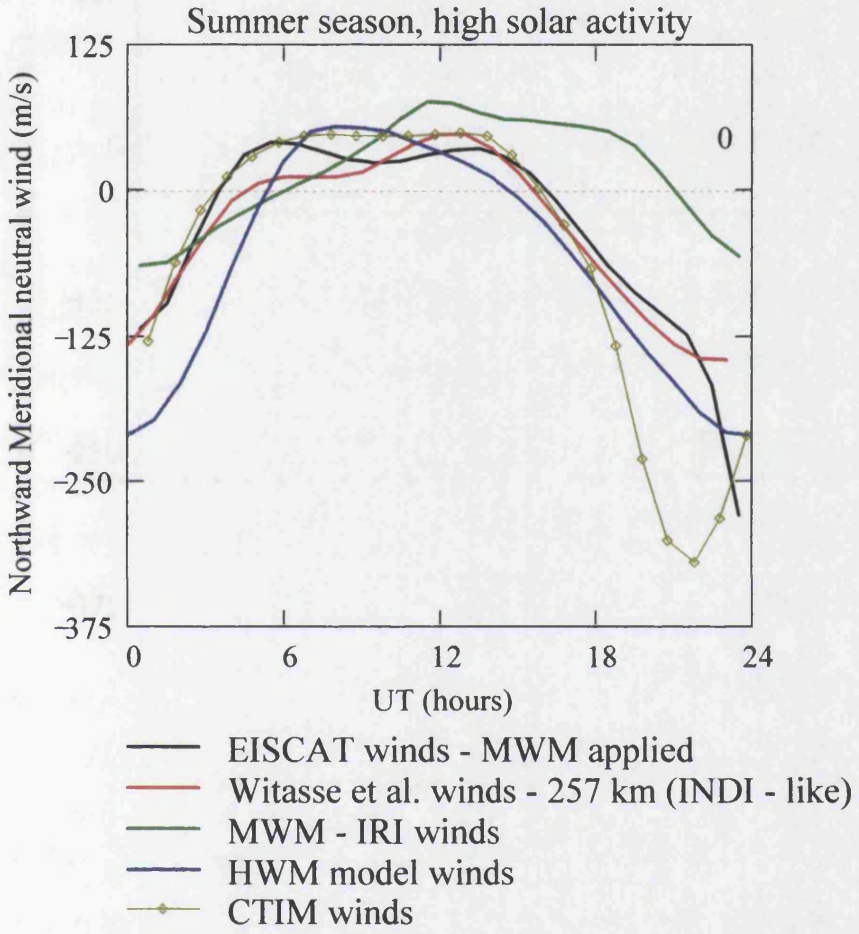


Figure 6.8 Combined contributions of summer season, high solar activity thermospheric neutral winds, positive northward, from each of the sources previously presented.

Repeating the trend seen for spring, the CTIM winds appear to reach their maximum southward winds during the night 1-2 hours earlier than all the other techniques and are the most southward winds from 1800 UT to near 2400 UT. In this season we see there is a much larger diurnal amplitude in the CTIM winds than for spring but also less evidence of the PNA than in spring. Examining the HWM winds reveals similar diurnal amplitudes from spring to summer but less influence of the PNA in the summer winds than in the spring winds. Finally the WLLP winds display

a much smaller diurnal amplitude in summer compared to spring but show similarly clear evidence of the PNA.

6.4.3 Autumn season

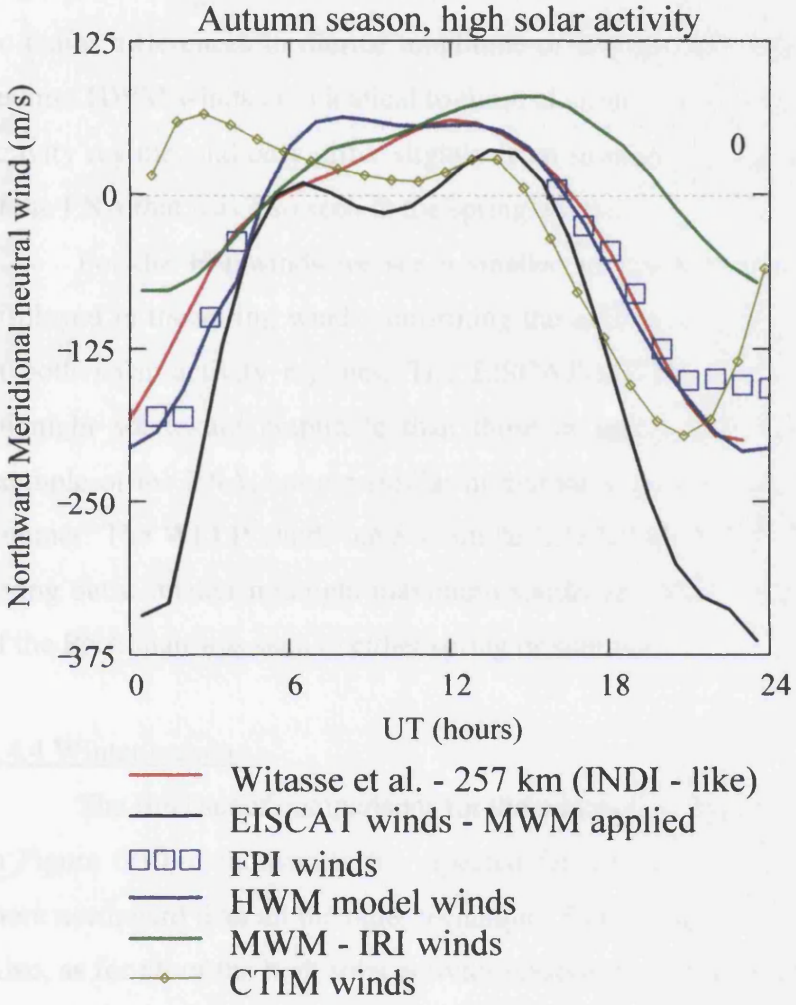


Figure 6.9 Combined contributions of autumn season, high solar activity thermospheric neutral winds, positive northward, from each of the sources previously presented.

For the autumn season at high solar activity results, presented in Figure 6.9, we see that in common with all the high activity seasons most techniques display pronounced evidence of the PNA. The IRI-MWM derived winds are again significantly more positive than all the other techniques from midday on to early in the morning and here most in disagreement with the EISCAT-MWM derived winds. EISCAT-MWM is the most equatorward of the techniques from 1800 UT to 0400 UT. Both HWM and the WLLP winds appear in good agreement with the FPI winds.

For these autumn winds CTIM is exactly the same as at spring due to the model formulation as was explained in the low solar activity case. CTIM produces the most poleward winds from around 2400 UT to 0600 UT and appears to reverse from southward to northward winds during the night earlier than any of the other techniques. The IRI-MWM winds are very similar to those at spring and summer with no major differences in diurnal amplitude or any greater evidence of the PNA. The autumn HWM winds are identical to those at spring, as was the case for the low solar activity regime, and only differ slightly from summer by displaying the larger degree of the PNA that was also seen in the spring winds.

For the FPI winds we see a smaller southward midnight amplitude than is displayed in the spring winds confirming the existence of the equinoctial asymmetry in both solar activity regimes. The EISCAT-MWM winds display a larger post-midnight southward amplitude than those at spring and also a more pronounced example of the PNA, but are similar in diurnal amplitude, being much larger than at summer. The WLLP winds have a similar diurnal amplitude in autumn compared to spring but a smaller midnight maximum southward wind. There is also less evidence of the PNA than was seen in either spring or summer.

6.4.4 Winter season

The final set of comparisons for the winter season at high solar activity, shown in Figure 6.10, demonstrates the repeated feature that IRI-MWM derived winds are more northward than all the other techniques from midday on to early in the morning. Also, as for all of the high solar activity seasons apart from summer, they are most in disagreement with the EISCAT-MWM derived winds, but there is no great change in the overall diurnal amplitude compared to the other seasons. There is clear evidence of the PNA in all the other techniques with full diurnal coverage. EISCAT-MWM is the most southward of the techniques from 1800 UT to 0400 UT. Both HWM and the WLLP winds appear in good agreement with the FPI winds until 0600 UT when there is a better fit between the FPI and EISCAT-MWM winds.

The CTIM winds have a smaller diurnal amplitude than is seen in any of the other seasons and also display a more pronounced PNA feature than any of the other seasons. CTIM is also the technique producing the most poleward winds from around 2300 UT to 0600 UT and the wind appears to reverse from southward to northward during the night around 3-6 hours earlier than any of the other techniques.

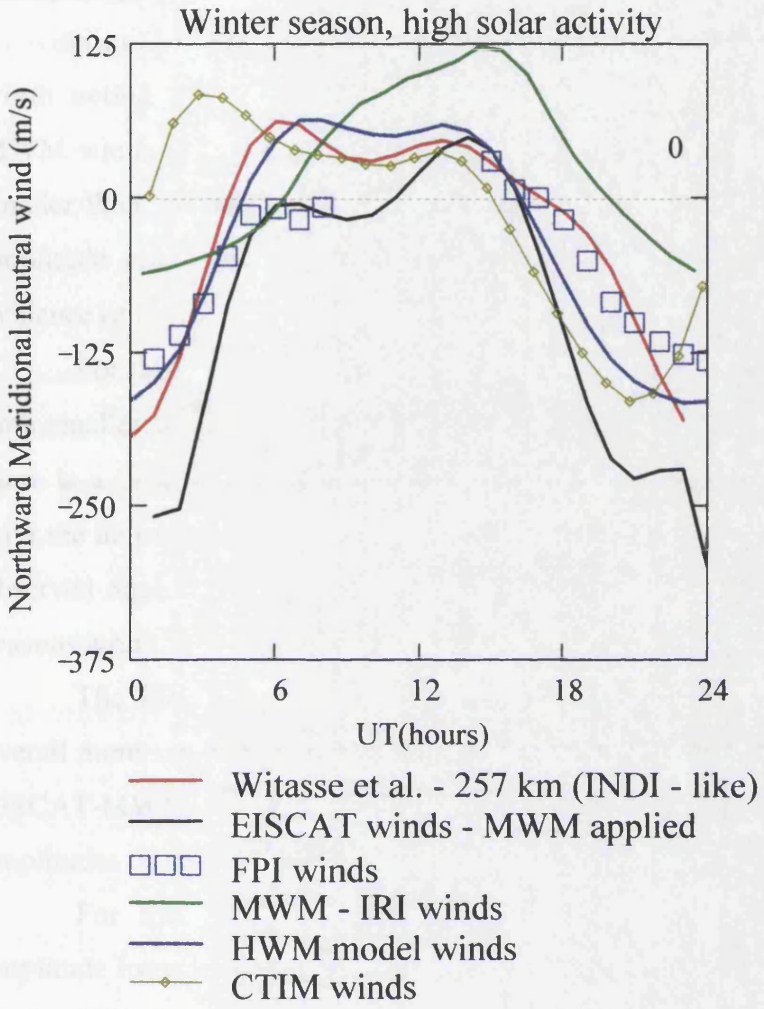


Figure 6.10 Combined contributions of winter season, high solar activity thermospheric neutral winds, positive northward, from each of the sources previously presented.

The FPI winds have smaller midnight southward winds than either of the equinox seasons. The HWM winds in winter display similar diurnal amplitudes to those in the other seasons and possibly the most prominent evidence of the PNA of any of the seasons. For the EISCAT-MWM winds there is a smaller diurnal amplitude than at equinox but similar to that at summer. The overall profile is however more southward than is the case in summer. Finally the WLLP winds have a higher diurnal amplitude than in summer but similar to that at equinox. Midnight southward winds are similar to those at autumn, being smaller than those in spring but greater than those in summer.

6.4.5 Summary of the high solar activity comparisons

Similar to the low solar activity case a number of clear trends and features are worth noting from the collected high solar activity climatologies. The EISCAT-MWM winds display similar diurnal amplitudes for summer and winter, these being smaller than the amplitudes for the equinox seasons. The spring winds have lower amplitude southward midnight winds than in the autumn winds and there is clear evidence of the PNA in all of the seasons.

For the WLLP winds summer and winter diurnal amplitudes are very similar, both smaller than the equinox amplitudes which are also similar to each other. Also there is a difference between the southward midnight maximum winds at equinox, with the autumn winds having less amplitude than those in spring, matching the FPI observed equinoctial asymmetry. There is some evidence of the PNA in all of the seasons with the most evidence possibly in the summer and winter seasons.

The FPI winds display the previously reported equinoctial asymmetry with overall more southward winds in spring compared to those in autumn, contrary to the EISCAT-MWM winds. The winter winds display the least southward midnight amplitudes of the three seasons for which data is available.

For the IRI-MWM winds there is a small trend of decreasing diurnal amplitude from winter through the equinoxes and into summer. There is also evidence of the PNA in all of the seasons with a small trend of increasing evidence from summer through the equinoxes to winter. There is no significant difference between the equinox seasons.

The HWM winds display consistent midnight southward amplitudes through all the seasons. Apart from the lower daytime northward winds in summer the diurnal amplitudes are also consistent. The evidence of the PNA however increases from summer through the equinoxes to winter. As with the low solar activity results no equinoctial differences are produced by the model.

For the CTIM winds we see that there is a consistent increase in the diurnal amplitude of the winds from the winter season through the equinoxes to the summer season which has the largest amplitude. The evidence of the PNA grows in the opposite direction with the least evidence in the summer winds and the winter season having the most evident abatement.

The high solar activity results can be examined in terms of the possible influence of the bias towards more geomagnetically active times at the equinoxes

compared to the solstices, as postulated by Lal (1996), as was done for the low solar activity case. Again, this would suggest a trend towards more southward diurnal mean winds at the equinoxes by comparison with the solstices. Of the measured winds the EISCAT-MWM and FPI winds appear to display evidence of this effect which was also found for the low solar activity case. However at high solar activity there is also evidence of this effect in the WLLP winds, which was not seen in the low solar activity case. While it has been noted in Section 5.2 that Buonsanto et al. (1997b) suggested a possible reason for the differences between winds derived from ISR measurements using the different techniques on individual days, it is difficult to see how this would produce a systematic difference in climatological results. None of the modelled winds shows clear evidence of the effect which matches the scenario at low solar activity.

In one of the earliest seasonal analyses of ISR derived winds Emery (1978) used solar maximum data at 43^{o}N to show a large seasonal change in mean southward wind from about +70 ms^{-1} in summer to –30 ms^{-1} in winter. On comparison with the data presented here it is interesting to note that the measurement based climatologies, i.e. WLLP and EISCAT-MWM, both exhibit the reverse trend, while all of the modelled winds match the trend. This may again indicate an over-reliance on mid-latitude results influencing the model output.

6.5 Intercomparison of the separate solar activity climatologies

The previous sections have shown the intercomparisons based on season within each of the solar activity regimes but here we will examine the comparisons between the seasons based on solar activity. Consistencies between the two sets of climatologies include:

1. The behaviour of the CTIM winds, which show the same trends of increasing diurnal amplitude from winter through equinox to summer and the increasing evidence of the PNA from summer through equinox to winter.
2. The HWM diurnal amplitudes are consistent through the seasons and have increased evidence of the PNA from summer through equinox to winter.

3. The FPI winds demonstrate smaller southward velocities at winter than for the equinoxes and the spring winds are more southward at midnight than those in autumn.
4. The IRI-MWM winds demonstrate the same diurnal amplitude trend, increasing from summer through equinox to winter.
5. The WLLP winds show greater midnight southward winds at spring than at autumn.

There are a number of clear differences also in the data:

1. In the IRI-MWM data there are greater midnight southward winds in spring compared to autumn for the low solar activity case whereas no such clear evidence is available for high solar activity.
2. For the HWM winds the winter season shows lower midnight southward winds than the other seasons at high solar activity, an effect which is not evident at low solar activity.
3. The EISCAT-MWM winds show clear evidence of the PNA in high solar activity for all seasons whereas no such evidence is found for the low solar activity case. Also in high solar activity the midnight southward winds are greater in autumn than in spring whereas the opposite is the case for the low solar activity case.
4. For the WLLP winds clear evidence of the PNA is seen in high solar activity for all seasons whereas no such evidence is found for the low solar activity case.

For low solar activity conditions we saw in Section 6.3.5 that large discrepancies occur between mid-latitude results and the high latitude results presented here. Igi et al. (1999) present mid-latitude results which demonstrate that the diurnal mean winds are most southward at high solar activity and most northward at low, a result which disagrees with the findings of Buonsanto (1991) also from mid-latitudes when no dependence on solar activity for mean diurnal winds was found. Comparing all of the seasons in each solar activity regime in the results presented here it emerges that the techniques show more northward winds in high solar activity than at low. This disagrees with both mid-latitude studies in the sense that it does reveal a

trend, contrary to Buonsanto (1991), but it is opposite to the trend found by Igi et al. (1999).

A study of the influence of solar activity on thermospheric winds was carried out by Hedin et al. (1994), using data from ISR, FPI, and ionosonde *hmF2* values at mid-latitudes. Low magnetic activity periods were examined for solar activity variations following the analysis scheme suggested by CAT (Hagan et al. 1993). They found diurnal mean wind trends with solar activity to be fairly weak except for Millstone Hill ISR which showed a shift in amplitude from strong southward to near zero or northward wind with increasing activity. This agrees with the IRI-MWM winds in the present study and the same trend is found to a lesser extent in the WLLP winds. The EISCAT-MWM winds however display the opposite trend and to the extent that a conclusion can be drawn as to the diurnal trend from the FPI data, without having full diurnal coverage, it would appear the FPI winds are also more southward for high solar activity.

Hedin et al. also show decreasing diurnal amplitudes with increasing solar activity during all seasons, except for Saint Santin ISR data which showed a slight increase in summer. Equivalent winds from *hmF2* data had strong decreases in diurnal amplitude in all seasons with increasing solar activity. The same trend was found by Miller et al. (1997) for ionosonde winds derived using MWM. By comparison in the present study we find the same trend only in the IRI-MWM winds with the EISCAT-MWM, FPI and WLLP winds all displaying the opposite trend, with larger diurnal amplitudes at higher solar activity.

Near midnight Hedin et al. found that most of the data also showed that the typically southward winds weaken with increasing solar activity in all seasons except summer, when results were mixed. Miller et al. (1997) found that looking at the global picture there was also a strong inverse dependence on solar activity as measured by the $F_{10.7}$ index for nighttime winds. However no dependence on $F_{10.7}$ was found for daytime winds. Again in comparison with the winds found in the present study only the IRI-MWM winds support this finding with EISCAT-MWM, FPI and WLLP winds all showing an opposite trend of greater southward winds at higher solar activity.

Duboin and Lafeuille (1992) in analysing ISR derived winds at Saint-Santin (mid-latitude) found that for both low and high solar activities summer winds are

more southward than winter winds. The maximum southward wind, near midnight, was however independent of solar flux within a given season. Also they find a shorter period of northward wind in the summer for low solar flux than for high, the same feature at equinox but no evidence of the same in winter. By comparison, the WLLP results shown here demonstrate the daytime northward wind periods are longer in high solar activity for all seasons, especially in summer. This is true also for the IRI-MWM winds but this reflects an overall shift in these winds to more northward values at high solar activity in all seasons. In all of the measurement based techniques the maximum midnight southward wind is seen to have a solar activity dependence, contradicting the findings of Duboin and Lafeuille. There is some evidence of the trend toward more southward winds in summer compared to winter in the HWM winds in both solar activity levels but no similar trend in any of the other techniques.

6.6 Conclusions

The previous sections have presented perhaps the most comprehensive comparison of techniques so far produced for the evaluation of high latitude thermospheric meridional neutral winds for a single site. Uniquely it allows comparison of direct measurements from the FPI with measurement based techniques in the form of the EISCAT WLLP and MWM climatologies as well as models which are both empirical in the case of HWM and IRI-MWM and theoretical in the case of CTIM. These climatological comparisons allow evaluation of the performance of the various techniques and sources without the influence of short term weather-like phenomena which affect more limited comparisons.

Presentation of climatological results for meridional neutral wind from ionospheric parameters at a single site without reference to any other techniques, such as Buonsanto (1990, 1991) and Igi et al. (1999) for mid-latitudes, is of limited use in establishing the influences that determine the gross features of the diurnal wind variation at that site. Few climatological studies include direct measurements from FPI for comparison with data derived in this way, although Dyson et al. (1997) include different *hmF2* techniques, HWM winds and FPI measurements for a study limited to a few days.

Here we can establish the reliability of the separate techniques based on derivation of winds from ionospheric parameters (EISCAT-MWM and WLLP) with external reference to the direct FPI winds, which has been lacking in previous studies

where there were either similar comparisons but no FPI winds (Buonsanto et al. 1997b) or FPI winds with these techniques limited to a short period of time (Buonsanto et al. 1997a). We can also test the abilities of the various models to follow the most important features in the diurnal wind pattern by comparison with the directly measured FPI winds and also contrast this performance with that of the remaining measurement based techniques.

The most important comparisons made here are probably those between the WLLP winds and the winds derived using MWM with EISCAT *hmF2* values. Both climatologies are derived from the same database of CP-1 experiments. The HWM is mostly derived from mid-latitude measurements so differences between HWM and other models are to be expected to a certain extent. Of course the importance of including the HWM climatologies is to establish how well they do compare and under what conditions the comparisons seem to be worse as this model tends to be the most often used for neutral winds when measurements are not available. The climatologies derived using MWM with IRI *hmF2* values provide another empirical model whose application and comparison allows evaluation of the applicability of the model (IRI) at a specific latitude. The CTIM climatologies provide a test of a purely theoretical model and as such the CTIM winds may be expected to have more systematic discrepancies compared with the other techniques due to the lack of any restraints applied based on actual measurements either at the site or extrapolated from elsewhere. It is important to identify and address these discrepancies however to establish how the model may be improved.

There is a great degree of similarity between the variations seen in the WLLP climatologies and those in the EISCAT-MWM climatologies by comparison with all of the model based climatologies. Unfortunately the lack of full diurnal coverage for the FPI data in any of the climatologies precludes comparisons in terms of the full diurnal variation with direct measurements. The major discrepancy seen from the plots of the individual climatologies is the fact that the IRI-MWM climatologies display greater midnight southward winds for low solar activity than for high activity. This is the opposite of the solar activity dependence seen in the EISCAT-MWM, FPI and WLLP winds. The dependence for the HWM and CTIM winds is less clear than for the other climatologies. It was pointed out in the comparison of the results presented here with those of Hedin et al. (1994), in the previous section, that the IRI-MWM results seem to exhibit the most mid-latitude like characteristics. In general the

results from IRI-MWM presented here have been seen to be the least realistic in terms of the trends displayed by comparison with all of the other techniques. This leads to the conclusion that the IRI is dominated by mid-latitude results and cannot properly represent the climatological variations in the *hmF2* values at this high latitude site.

At high solar activity as has been shown earlier there is more evidence of the PNA for both WLLP and EISCAT-MWM winds. The fact that this is a feature seen in both the EISCAT derived datasets indicates it is a real effect which is not being repeated properly by the models. Unfortunately it is not possible to test this conclusion properly with the FPI data due to the lack of full diurnal coverage. For mid-latitude sites the most likely cause of the PNA, as evidenced by the corresponding semidiurnal components, has been identified as the influence of the semi-diurnal tide (Buonsanto, 1991; Hagan, 1993). CTIM simulations suggest that in-situ semidiurnal oscillations generated at mid- to high latitudes by ion-neutral interactions such as ion drag and Joule heating are comparable in magnitude and, towards higher latitudes, stronger than the upwards propagating tides with which they interact through both destructive and constructive interference (Muller-Wodarg et al., 1999). The results here allow these CTIM predictions to be tested with measured winds and the clear solar activity dependence of the evidence of the PNA, and hence of a semidiurnal component in the winds, will help direct the efforts to investigate the balance between the possible sources for these components.

Previously, WLLP compared their results with the Kiruna FPI wind data presented by Aruliah et al. (1996a). The FPI dataset in that study included the majority of the data used in the present study for the derivation of FPI climatologies. As described earlier Aruliah et al. showed that the March equinox nighttime winds are larger than the September equinox winds during high and low solar activities and that this asymmetry is greater for high solar flux. These features have been demonstrated in the FPI climatology presented earlier in this chapter. These features are repeated in the WLLP dataset and were noted as such in their study. WLLP found large discrepancies however in the midnight winds between their technique and the FPI winds as they find their amplitudes can be almost a factor of two larger. In attempting to resolve the differences between the two datasets WLLP point to the weaknesses of FPI measurements, discussed in Aruliah et al. (1996a), and the small uncertainties of their method. By implication this claims their results more accurately reflect the true nature of the wind variations.

With regard to these differences however WLLP ignore the findings of their own comparison with a previous study using direct wind measurements. Fauliot et al (1993) (FTH) presented a model of the F-region horizontal wind at high latitude from the MICADO interferometer measurements during three winter campaigns from 1988 to 1991. This model gives meridional winds in winter for three magnetic activities and high solar fluxes at a typical height of 250 km for typical times between 1400 and 0500 UT. WLLP made comparisons between the FTH model (specifically that corresponding to their medium magnetic activity regime) and their winter season, high solar flux data. At midnight the FTH model underestimates the wind velocity by comparison with the WLLP winds by around 40 ms^{-1}. It can be seen in the full set of comparisons for this season presented in Section 6.3.8 that this brings the FTH winds into better agreement with the FPI winds than those of WLLP, demonstrating consistency between the two sets of direct measurements. The small database used for the FTH model in comparison to those presented in this study makes it impossible unfortunately to compare the results for any of the other seasons or at different solar activity levels.

The findings of this comparison with FTH are probably a more pertinent test of the accuracy of the wind measurements from FPI and WLLP than individual assessment of the relative accuracy of the techniques as many of the factors involved are subjective. WLLP have previously suggested that their climatology from the EISCAT CP-1 data be included in any future re-formulation of the HWM. The results of the comparisons here indicate that for the seasons and UT coverage provided by the FPI, the FPI results should be used. However, for times and seasons not covered by the FPI, given the lack of data coverage in the HWM, that the WLLP winds would provide useful completion of the diurnal variation and should be included.

The assertion by WLLP that geomagnetic activity, as measured by A_p index value, cannot be regarded as a pertinent separation criterion for analysis of climatological effects, alongside season and solar activity level, is quite simplistic. In Section 5.2 the influence of geomagnetic history was displayed when examining individual nights with similar current 3-hour A_p values, and explained in terms of the series of 3-hourly values preceding the displayed data. In Section 6.2 we saw that the difference between the WLLP study and those undertaken previously was attributable to the use of contemporaneous local measurements to filter out disturbed times. This

however isn't an argument rejecting geomagnetic activity as a valid separation criterion in climatological studies but rather an indication of the limitations of the use of the standard indices, especially the mean daily values.

The problem is noted by other authors. Biondi et al (1999) used the "current" K_p values to select geomagnetically quiet conditions. However, they acknowledge the nature of the equatorial ionosphere is strongly dependent on the history of the geomagnetic activity for up to 30 hours earlier, as pointed out in numerous studies (e.g. Fejer and Scherliess, 1997). Since use of this more stringent criterion would have significantly decreased the number of wind measurements that could be used in their database and since there is no a priori means of determining the length of this period for a given observation night, they adopted the simpler "current" criterion. When using the current criteria with FPI results it is important to remember that these are direct measurements from the neutral gas. The techniques that use ionospheric measurements to derive neutral winds can be expected to be more sensitive to the geomagnetic conditions and therefore to require more stringent filtering out of geomagnetically active conditions. This is due to the ionospheric parameters, such as *hmF2*, changing under the influence of enhanced activity much more rapidly than changes which occur in the thermosphere. This has an effect whereby the derived winds in these disturbed conditions reflect ionospheric variability rather than thermospheric variability. It is probable in this case that using local measurements of the electric field for instance is a necessary precaution before using data derived using the ISR techniques.

The CTIM winds have been investigated to address the possible reasons for the consistent differences between CTIM winds and winds from other techniques. One factor examined was the influence of the geomagnetic activity level used in generating the CTIM output. A comparison showing the CTIM diurnal wind pattern for different K_p levels is shown in Figure 6.11. The model output is for high solar activity conditions ($F_{10.7}$ = 200) in summer. The diurnal wind patterns from this model output for Kiruna are shown for three different levels of K_p, with the highest K_p level (K_p = 5) giving the greatest evidence of an early daytime poleward peak. There is also a trend toward larger southward winds at night reaching their peak earlier, at higher K_p. Full climatologies in terms of season and solar activity level are not available for all these K_p levels from the standard model output. However for

specific conditions, such as $K_p = 2^+$, comparative model output for all conditions is available.

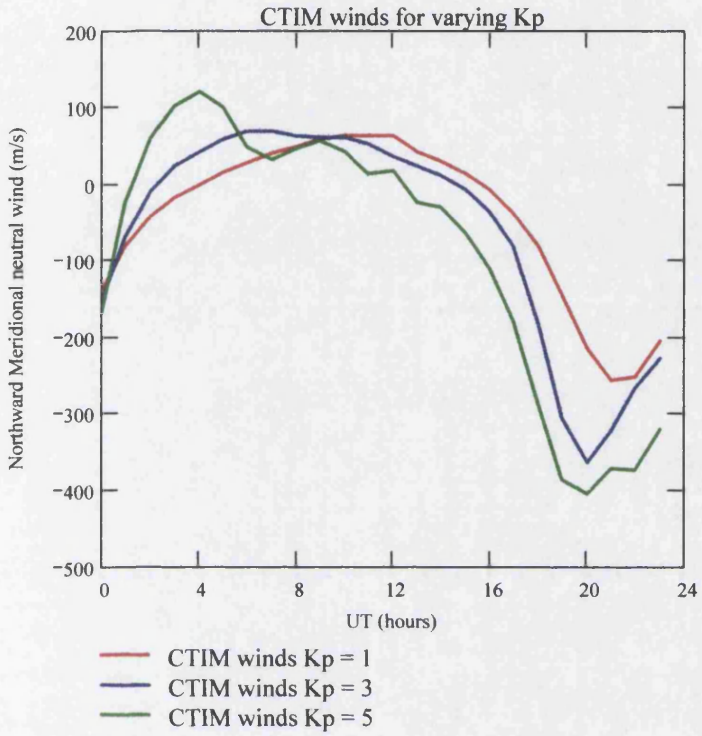


Figure 6.11 Thermospheric meridional neutral winds from the CTIM model for three different K_p levels.

When the level of geomagnetic activity is set in CTIM, as determined by the K_p value, a constant average electric field model is applied to the coupled ionosphere and thermosphere coded in the model. However, the real electric field has rapid, random fluctuations that are smoothed out in the electric field models. This leads to the possibility that the influence of the electric field may be overestimated even in relatively quiet scenarios. To test this aspect of the way the model works Figure 6.12 plots the ion drag force values through the day for the same K_p levels plotted in Figure 6.11, while Figure 6.13 plots the corresponding pressure gradient force values, derived from the same model output. Comparing the ion drag values to the winds shown in Figure 6.11 we can clearly see that while the winds are highly modified by

the influence of the ion drag during the 2000-2400 UT period, it does not explain the evidence of large northward winds in the early part of the day at high K_p.

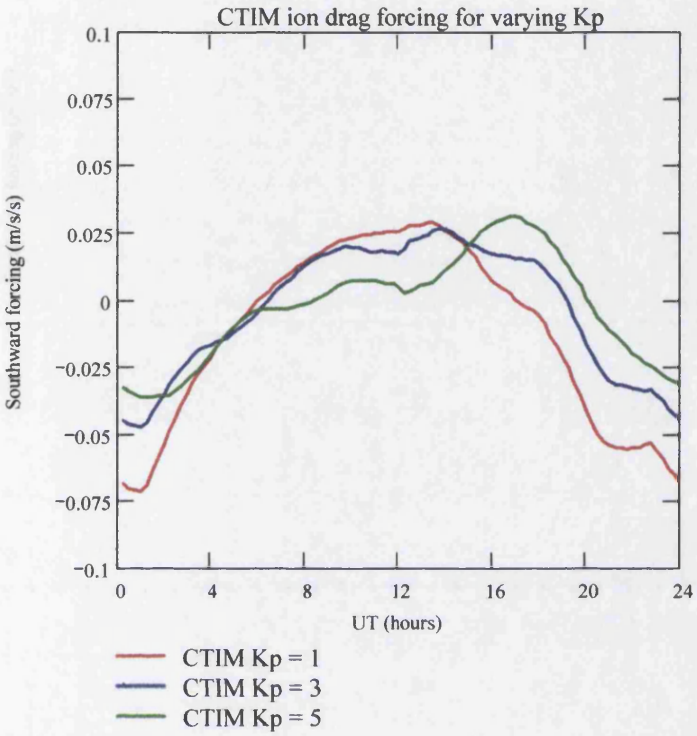


Figure 6.12 CTIM ion drag values at different K_p levels for Kiruna.

In Figure 6.13 the corresponding values of pressure forcing are shown and in conjunction with the values for ion drag it is possible to see the reasons for the alteration of the winds early in the day in reaction to the changes in K_p level. The pressure forcing changes greatly in nature between the K_p levels with much more variation through the day at the higher K_p levels, demonstrating the complex balance at this location between pure solar heating driven forcing, as is seen in the $K_p = 1$ case, and the influence of the energy input at higher latitudes, as it is represented in the model. The combination of these forces, together with smaller contributions from the other forcing terms, produces the resultant winds shown in Figure 6.11 and demonstrates the sensitivity of the CTIM output to the chosen K_p level.

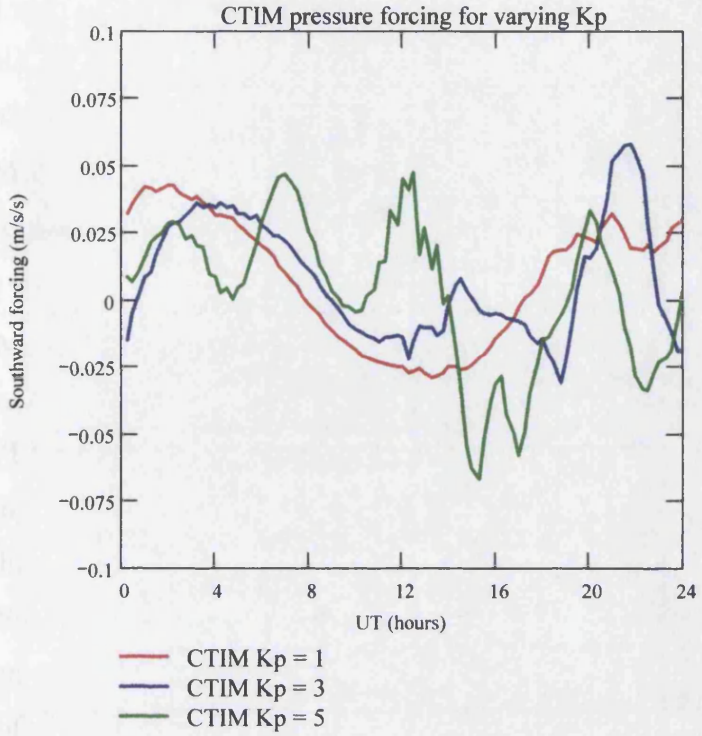


Figure 6.13 CTIM pressure forcing values at different K_p levels for Kiruna.

Another possible influence of the sensitivity to K_p in CTIM is the apparent phase shift in the CTIM winds by comparison with all of the other techniques in most seasons. This is evident in northward winds being seen earlier in the day and the midnight maximum also being seen earlier. This may be due to a combination of the over-sensitivity of the momentum forcing terms, such as the ion drag and pressure shown here, to the K_p level. This would then produce a balance in the forces which is unrepresentative of the actual balance at this location, resulting in winds which are shifted by respect to the actual winds.

The ability to extract the forcing terms from CTIM to examine the reasons for the characteristics in the output winds allows the fundamental physics of the thermosphere as represented in the model to be tested. In this manner the ways in which the model may be improved can be addressed. For example, the representation of the high latitude energy inputs and also the ion drag calculations will be addressed by taking into account the work of Codrescu et al. (2000), who have parameterized

the random electric field fluctuations to increase the Joule heating appropriately. Updated model output will then be compared to the measured climatologies presented here. This should produce a more realistic representation of the changes in the high latitude thermosphere to the level of geomagnetic activity.

Miller et al. (1997) found that in making comparisons between HWM winds and those derived using MWM with ionosonde data that the HWM winds were nearly constant with $F_{10.7}$ except at high magnetic latitudes. We see some variation between the HWM winds at high and low activities here but not to the same extent as is seen in the majority of the other climatologies, including all of the experimentally derived ones (WLLP, EISCAT-MWM, FPI). The HWM results can be explained in this context by reference to the data which has been included in the model formulation. For the high latitude European longitudinal sector the only contributing winds come from EISCAT and an FPI at Svalbard. However the EISCAT data used only covers the region from 100 to 120 km in the lower thermosphere. The Svalbard data is from periods of moderate to high solar activity and only for the winter season.

The fact that the HWM winds are seen to resemble the high solar activity winds in the low solar activity examples points to a lack of coverage of low solar activity conditions in the dataset. Further, the strongest seasonal response is seen in winter at high solar activity when the results are closest to this contributing dataset. There is also clear evidence that the variations in diurnal amplitudes produced by HWM are not large enough to match those found in the results from the measurement-based techniques. Khachikjan et al. (1997) found HWM winds even at mid-latitude to produce diurnal amplitude variations which were too small in comparison to winds derived using IRI model results.

Miller et al. (1997) also suggest that while the latitude dependence of the HWM winds and those derived from ionosonde data are similar, there are large differences between the actual wind values between the two techniques which could be caused by the dominance of the American sector in the HWM dataset. They also find that large differences were found between HWM winds and those derived from ionosonde winds at high magnetic latitude in both hemispheres. For daytime winds at noon there was no significant difference between solar flux levels for the winds derived from ionosonde data. For the same noon conditions HWM showed smaller winds at solar minimum. For the IRI-MWM climatologies presented here we see the

same trend to smaller daytime amplitudes at low solar activity but no similar evidence in any of the measured climatologies or in the theoretical CTIM climatologies.

WLLP found very good agreement with HWM around midnight, except in the summer and high solar flux case, and the autumn and low solar flux case, where the difference can reach around 50 ms^{-1}. For solar maximum periods relatively good agreement has been found between 1500 and 0500 UT, except in summer in the morning and in winter between 1600 and 2200 UT. During the daytime HWM generally overestimates the northward wind (around 30 ms^{-1} in summer at 0800 UT) by comparison with the WLLP winds. For solar minimum periods, larger discrepancies appear with the HWM model overestimating the wind between 0500 and 1000 UT and underestimating it by around 60 ms^{-1} in winter at 1200 UT. Comparisons to the other techniques presented here reveal broad agreement with these conclusions. The general overestimation of daytime winds until around 1200 UT is most evident in the low solar activity cases.

In justifying the differences between their winds and the FPI winds (see earlier in this section) WLLP also note that their midnight winds are in very good agreement with the HWM model at 250 km. The comparison with HWM winds may be less than ideal in terms of the relative independence of the techniques also as it is constructed mainly from ISR data at mid-latitude. It should also be noted however that the FPI winds are direct measurements and do not invoke as many assumptions as are made in the derivation of the neutral winds from ISR data. Only when more data with better coverage of seasonal and solar activity conditions from this latitude have been included in the HWM will its results be useful in terms of determining the validity of measured datasets.

It is pertinent to address the findings of Buonsanto et al. (1997b) from a mid-latitude study, and demonstrated by Figure 5.8, that the winds derived from ISR derived *hmF2* values, as here for the EISCAT-MWM values, were too northward at midnight by comparison with those from the corresponding WLLP-type winds. We have seen in the discussion of the comparison between solar activity levels that there is a clear influence of the increased geomagnetic activity at equinox, as postulated by Lal (1996). This is most evident in the EISCAT-MWM winds but the WLLP winds don't seem to react in this manner. This indicates that there may be a tendency for the EISCAT-MWM technique to be more sensitive to geomagnetic activity than is the

case for the WLLP winds. It may be that the use of the FLIP model output for the "balance height" *hmF2* values generates this difference as these values are a critical parameter in the determination of the EISCAT-MWM neutral winds. While they will change to reflect the A_p value input to the MWM they cannot reflect rapid changes in geomagnetic conditions through the course of an experiment.

A possible source of systematic discrepancies when using the MWM, as with the EISCAT-MWM and IRI-MWM winds here, would be provided if there was a problem in the application of the FLIP model at this latitude. The intended boundary for the MWM is from 15^o to 70^o in geomagnetic latitude and so at Tromso we are right at the edge of the intended application of the model. The balance heights chosen by the MWM, based on the parameterised FLIP output and the various input parameters, may be considered as the most likely factor to suffer from errors in this regard. However examination of the climatologies derived using the MWM from two different sources, i.e. the EISCAT *hmF2* and IRI *hmF2* values, shows large differences both on an individual season and activity level (Figures 6.3 – 6.10) and also on the climatologies taken as a whole (Figures 5.7 and 5.15). These differences are evident in both amplitude and shape and lead to the conclusion that any influence of the FLIP model outputs in shaping the climatologies artificially is much less than the difference arising from the use of different input datasets. Hence there is no evidence from the climatologies presented here (EISCAT-MWM and IRI-MWM) that the use of the FLIP model introduces any systematic influence on the derived winds when using the MWM.

According to Titheridge (1993, 1995a) neutral winds derived using servo theory implementations are inaccurate during the sunrise and morning period because of a shift in the zero-wind F-region peak downward from the balance height. This would lead to an overestimate in the MWM derived northward winds during this period as the observed *hmF2* would be interpreted as being due to the influence of the neutral wind alone. Miller et al. (1997) however believe that this problem will be compensated in the MWM by the photo-chemistry that is contained within the FLIP model. In examining the evidence for this we see that in both summer season climatologies the EISCAT-MWM derived winds are indeed more northward than the WLLP derived winds. However there is no clear evidence of a similar trend in the autumn and spring climatologies. Even more striking however is the fact that both the

winter climatologies show the reverse trend with the WLLP winds being more northward than those from the EISCAT-MWM values. Interestingly however the IRI-MWM derived winds do show a consistent post-sunrise increase in northward winds to higher northward values than are evident in the WLLP winds. This is true for all but the winter season, high solar activity case for which the winds in this period are similar. This indicates that while the IRI modelled *hmF2* values seem to suffer from the problems identified by Titheridge when deriving neutral winds, the effects are not apparent when using measured values of *hmF2*.

6.7 Further work

In general the development of models of the thermospheric meridional neutral wind has been assisted through recent years by comparison with direct measurements from FPIs and winds derived from radars and ionosondes (Buonsanto et al., 1989; Hedin et al., 1994; Buonsanto et al. 1997b; Dyson et al., 1997; WLLP). This process will continue as the database of measurements is extended and the gaps in the data coverage are addressed, leading to more detailed and specific comparisons being undertaken. Therefore it is envisaged that updates of models as they come along will be tested by reference to the measured climatologies as was done here. While examples of both empirical and theoretical models are used here for the comparisons there are other models such as NCAR TIEGCM and VSH which could be used in further comparisons. Some specific examples of the work possible to continue and extend the work presented here follow in the next few subsections.

6.7.1 FPI's

The FPI database at Kiruna has been demonstrated to be of particular use in this work as it allows climatologies from other methods to be compared to direct measurements. The problem with FPI data however is that at present observations are restricted to dark, nighttime conditions which at Kiruna means there is no data for the summer season and a gap corresponding to daytime conditions in all of the other seasons. The development of an FPI with daytime observation capabilities would allow comparisons with direct measurements at all times of the day and during all seasons. A number of attempts have been made to get this type of instrument into regular operation (e.g. Narayanan et al., 1989) and it would obviously take time to develop climatologies for the extended FPI dataset provided. In recent years

experience in the APL with double etalon Fabry-Perot devices and bare CCD detectors has allowed the possibility of such an instrument being developed to be suggested and it is hoped that examples will be tested in the near future.

Another FPI has been in operation at Skibotn, close to Tromso in Norway since 1996 (Kosch et al., 1997). While there is only a limited amount of data currently collected by this instrument, collaborative efforts are underway and the data may be used in the future to test whether the assumptions made in the comparisons presented here in terms of horizontal scale size are valid.

The FPIs operated by the APL are constantly being improved and updated and it is expected that current detector and software development will lead to better time resolution for all of the current field instruments. This will allow better co-ordination between the time resolution available for the FPI and that of the EISCAT radar, the new ESR radar and other instrumentation, such as the ionosondes, which might be used in further co-ordinated studies of the ionosphere-thermosphere system. The latest design for a wide-field, scanned etalon FPI coupled to an intensified CCD system is known as SCANDI and will hopefully be deployed in the near future.

As all of the APL FPI's develop better seasonal and solar activity coverage within their specific datasets, climatological comparisons may be undertaken to examine different influences more closely. The effects of, for instance, geomagnetic activity and IMF conditions might be envisaged as has been done for shorter periods, examining weather-like phenomena (e.g. Thayer et al. (1995)).

One of the modelled parameters used in the servo theory neutral wind derivation as implemented by MWM is the neutral temperature, which is determined by the MSIS model. It is possible to use the FPIs to measure this parameter, however this has only recently become a part of the standard operation mode due to the difficulties involved in providing accurate calibrations at the remote observation sites. Therefore in combination with the EISCAT radar it is possible to envisage using measurements of *hmF2* and ion and electron temperature from the radar with measurements of neutral temperature to calculate neutral wind "by hand" through the servo equations. These values could then be compared to the MWM derived values using the same *hmF2* values, EISCAT winds using the INDI method and the neutral winds as measured by the FPI. In certain conditions neutral temperatures may also be derived from the EISCAT results, which would allow comparisons to be made to the

FPI results and also provide daytime estimates to be used in the calculations when no FPI data was available.

Jarvis et al. (1998) point out that in order to put the observed long-term changes in *hmF2* observed by ionosondes in proper context, trends in thermospheric neutral temperatures also need to be examined. This places great importance on the combined measurements between colocated FPI's measuring the neutral temperatures and ionosondes measuring the *hmF2* values. At each of the APL FPI sites these comparisons are possible with the Svalbard CADI, Kiruna Digisonde and RAL Digisonde.

6.7.2 High latitude climatology

There are a number of non-FPI specific improvements that could be made for the high latitude climatologies and their inter-comparison. If enough data was available then it would be possible to take separate subsections of each climatology for comparison, e.g. taking data from the first low solar activity period contributing to the comparisons (1984-1987) and comparing it to the data from the second low activity contributing period (1993-1996). This has been investigated for the mid-latitude climatology from Millstone Hill and found to produce significant differences (Buonsanto and Witasse, 1999). In order to facilitate this type of investigation it will be necessary to extend the use of *hmF2* from EISCAT data to experiments other than CP-1. It is also possible to make use of further CP-1 experiments as they are performed and also to obtain data that is CP-1-like but run as special program (SP) experiments at EISCAT. Some of the data which has contributed to WLLP for instance is taken from special program data. Extending the database of measurements available for inclusion in the climatologies in this manner should also eliminate the possibility of non-uniform weighting in the data subsets. As mentioned previously the EISCAT data used in this study has been in the NCAR pre-analysed format. It is possible that improvements to the accuracy of *hmF2* determination could be made if the data were re-analysed with the *hmF2* determination as the primary consideration.

A recent comparison between electric fields derived from the Tromso Dynasonde and those derived from EISCAT measurements (Sedgemore et al., 1998) has shown the possibility of using the Dynasonde to produce the electric field correction to the neutral winds derived using MWM. This would mean that combined with the *hmF2* values derived from the Dynasonde the derived neutral winds would

be equivalent to those produced using the EISCAT radar. The co-location of the instruments would then allow intercomparison of the two techniques. This combination could test the possibility of using the Dynasonde in a manner which would allow it to do a comparable job to the radar in terms of the neutral wind determination by comparison with both the servo theory technique and the more usual technique by using the ion velocity measurements. This would provide a relatively cheap way of making reasonable neutral wind measurements, including daytime measurements, from these ionospheric instruments, which are orders of magnitude cheaper than an incoherent scatter radar.

While the winds derived from Kiruna Digisonde measurements have not been presented here in a climatological sense it would be possible to use the database of values recorded at Kiruna to derive a similar climatology to those presented here. It is possible for the Digisonde to operate in "ion drift" mode as described in the earlier chapters. This mode allows the electric field to be determined using the digisonde but this is a more intensive operating mode both in terms of post-analysis and instrumental wear and is not used as the standard operating mode. The data in the Kiruna digisonde database has not been recorded in this mode for instance. The lack of an available electric field correction for neutral winds determined by this instrument however would therefore lead to a need for caution in interpreting the results or alternatively in filtering the input data to the MWM. It might be possible to use some EISCAT electric field values from Tromso to correct the Kiruna Digisonde derived winds but this makes assumptions about the horizontal scale sizes that might be valid for neutral winds but probably not for electric fields. It would however allow some measure to be made of the influence of the electric fields as individual days of data appear to show very good agreement between the Kiruna Digisonde derived neutral winds and those from the EISCAT radar. In the future it might be possible to make electric field determinations with the Kiruna Digisonde used in drift mode, but this would still require long term operation and data collection before any climatological data could be retrieved for neutral wind determination.

There also exist a pair of ionosondes at lower latitudes in Sweden at Uppsala and Lycksele, which could also be used for neutral wind determinations to investigate the possible changes with latitude. These data could be used in comparisons with the RAL dataset which is more easily identified as mid-latitude data. In this case it would

be interesting to check against the predictions of the HWM, IRI-MWM and CTIM models again to see which produces the most accurate fit.

In the latest version of MWM as used in this study the value for the $O\text{-}O^+$ collision frequency has been set to the "Burnside factor" of 1.7 as adopted as an interim standard by CEDAR, through the use of the FLIP model which has this value "hard-wired". It should be emphasised that the motivation for choosing an interim standard was to facilitate intercomparisons between different ISR wind and other aeronomical measurements which had been unnecessarily complicated by different choices for the $O\text{-}O^+$ collision frequency. Some recent work has suggested that this correction value is too high and should be either changed to a lower value of 1.2–1.3 (Davis et al., 1995, Buonsanto et al., 1997c) or simply ignored by setting the value to 1.0 (Lathuilliere et al., 1997), with the most recent work pointing towards a value of 0.98 (Omidvar et al., 1998). There exists the possibility with the extended datasets prepared for this study to use a lower value for the Burnside factor in the FLIP model and use the revised parameterised outputs as the basis for a revised MWM which would then produce revised winds. In this manner an estimate of the improvement or otherwise could be evaluated both on an individual nightly basis and in a climatological sense. While Buonsanto et al. (1997b) have carried out such a study at mid-latitudes they used a less extensive dataset and provided less opportunity for comparison with other techniques. As pointed out in Section 5.2.2, they also use data from nights with geomagnetically disturbed conditions. While they may have considered this not to have been as much of a problem at mid-latitudes as at high latitudes, it would be preferable to have an extensive enough dataset to allow these periods to be excluded.

6.7.3 RAL

In the long term it will be possible to develop a similar set of climatologies for comparison at RAL to that at Tromso/Kiruna, from the joint measurements of winds from FPI and derived from the Digisonde measurements. The fact that the Digisonde routinely takes round the clock measurements that can be used in this context and that the database at RAL/Slough extends back to the 1930's should allow a very reliable climatology to be produced. However it will take some time to develop the FPI

climatology as the length of a solar cycle is unfortunately not a variable under our control.

At present tests are being carried out also to use the digisonde in drift mode to make measurements of the electric field to see if this is important in correction of the winds at this latitude (Davis, 1999). A significant electric field influence at mid-latitude under disturbed conditions would lead to the need for corrections to be made to MWM derived winds in a similar manner to those carried out for the winds derived from the EISCAT data.

The Buonsanto et al. (1997c) study, mentioned previously in Section 6.7.2, used collocated ISR and FPI measurements at mid-latitude to attempt a calculation of the $O\text{-}O^+$ collision frequency in an experiment similar to the EISCAT INDI experiment which was based on the early work of Burnside et al. (1987). As was suggested for the high latitude Tromso/Kiruna location a reworking of the MWM with different values for the Burnside factor could be applied to the data gathered at RAL from the Digisonde and compared to the FPI data. This would act as an independent test for the correct Burnside factor in conjunction with a similar analysis of the high latitude data. With the digisonde running in drift mode any correction necessary for the E field would be taken into account and a dataset could be built up to enable the comparisons to be made.

The longevity of the RAL/Slough ionosonde database leads to the possibility of taking different subsets of the data corresponding to each of the solar cycles included in the coverage and comparing them to see if they are consistent. If this confirmed the consistency of the results then the entire database could be used to create a very accurate climatology. There is also the possibility of using CTIM and HWM at this latitude for comparison and also using MWM with IRI *hmF2* values to test the trends we see at high latitude to see if they are repeated at mid-latitude, in terms of the intercomparison of the techniques. Comparison would also be useful with the many previous mid-latitude climatological studies.

With the FPI available to produce neutral temperatures there may be some value in producing servo equations "by hand" at RAL in studying shorter term effects in the mid-latitude thermosphere. In this type of study it would also be possible to use manual hand-scaling of the Digisonde data for better approximation of the *hmF2* values.

The recent study by Igi et al. (1999) reported results of climatological analysis of meridional neutral winds for both seasonal and solar activity effects from a database of ionosonde measurements made at Kokubunji in Japan. The authors note the lack of electric field correction in their results but say that restricting their data to times of low geomagnetic activity should allow them to have confidence in the results at this latitude. They note a number of differences in their results by comparison with previous studies such as by Buonsanto (1990,1991). This points to a longitudinal difference in results not easily explained at the moment. At RAL the same analysis may be carried out and also a comparison to FPI results to see if there is any noticeable systematic difference between the winds. Miller et al. (1997) compare global output of HWM and IRI *hmF2* winds which is important in resolving the longitudinal differences pointed out by Igi et al. (1999), by comparison with other similar mid-latitude studies. However this only compares two empirical models, whereas a comparison to a full TGCM model such as NCAR-TIEGCM or CTIM would allow better identification of the possible influences in the global sense by examining the relative importance of the forcing terms at different locations.

6.7.4 Svalbard

At polar cap latitudes such as at Svalbard the assumptions of servo theory are no longer valid as the angle between the geomagnetic field and the vertical is so small that neutral winds no longer provide a strong perturbing effect on the ionosphere and the component of the *hmF2* change resolved to this angle is too small to allow MWM to be applied.

There has been previous work done on neutral wind climatologies in the polar cap, such as Killeen et al. (1995) who present a unique database of geomagnetic polar cap FPI measurements which extends over the complete range of solar activity. The measurements from a total of 902 nights of observations were compared with the predictions of the VSH and HWM models. The APL has an FPI at Svalbard to provide direct measurements of neutral winds for which 6 years of observations have so far been gathered. Together with the output from both the HWM and CTIM models there is plenty of scope to investigate the influences on the neutral wind at this latitude both on an experiment by experiment and climatological basis. The recent construction of the EISCAT Svalbard radar (ESR) allows similarly detailed

investigations and comparisons to be undertaken between the Svalbard FPI and ESR as have been produced in the past between EISCAT and the Kiruna FPI.

Importantly Svalbard usually lies under the cusp region in daytime. This means the atmosphere above Svalbard can be expected to see the energetic thermospheric effects of magnetosphere to ionosphere coupling. Measuring the thermospheric temperatures here in common with the winds is very important as both measurements can be influenced by the particle precipitation and Joule heating in this region. The temperatures reflect most of the energy transfer to the thermosphere and are being derived from the FPI observations at present. Model studies such as that by Namgaladze (1996) show that it is possible for energetic input in the polar cap to transmit thermospheric variations to much lower latitudes with significant amplitudes. It is suggested that these are seen in both temperatures and winds and the importance of contemporaneous measurements from different latitudes of these parameters is emphasised. With the RAL, Kiruna and Svalbard FPI's representing mid-, high and polar latitudes the development and propagation of these influences and related phenomena such as travelling atmospheric disturbances (TADs) may be examined in detail.

6.8 Summary and conclusions

This thesis has presented the results of comparisons between the methods possible for deriving neutral winds from ionospheric measurements, direct measurements with an FPI, together with empirical and theoretical models as applied to the particular Kiruna/Tromso high latitude location. The many possibilities existing to improve and extend the work presented here have been discussed.

Some important results have been found regarding the comparison of these techniques. There are clear differences between the solar activity dependences of diurnal amplitude and mean and also the amplitude of nighttime equatorward winds as derived by the various techniques. The evidence for higher semidiurnal components in the winds at high solar activity than at low solar activity is seen in both WLLP and EISCAT-MWM winds and will be used to investigate the sources of these components, at this high latitude site, in conjunction with the CTIM model.

The HWM model, which is the most often used when measured neutral winds are unavailable, has been shown to lack sensitivity to both the seasonal and solar activity conditions at high latitude. In order to address this failing it is apparent that

further revisions of this model should include contributions both from the FPI database of direct measurements and that available from the EISCAT thermospheric measurements to complete the diurnal and seasonal coverage.

The CTIM winds presented here from standard runs of the model have shown systematic differences to the measured techniques. While the model output is not intended to be a simulation of the upper atmosphere in terms of matching all empirical measurements in all conditions, the differences in the trends with season and solar activity need to be addressed. One of the problems contributing to these differences has been identified as the use of a constant K_p index introducing systematic over-estimates in some of the modelled forcing terms. When these issues are addressed, by for instance the inclusion of new electric field implementations as suggested by the work of Codrescu et al. (2000), the results may be tested again by reference to the other results presented here.

The IRI-MWM technique agrees well with previous studies at mid-latitudes by Hedin et al. (1994) and the more global study of Miller et al. (1997) but these findings are contradicted by all of the climatologies based on measurements (WLLP, EISCAT-MWM and FPI), excepting the diurnal mean where WLLP also agree with the IRI-MWM. This leads to the conclusion that this model is overly dominated by mid-latitude measurements and cannot be used as a reliable source of *hmF2* from which to derive winds for this latitude.

Appendix A - Theory of the Fabry-Perot Interferometer

A.1 Introduction

The theory and principles of the application of the Fabry-Perot Interferometer (FPI) instrument are presented in this Appendix. Due to the unique properties of the FPI, having the highest luminosity of those spectroscopic devices that depend on the interference of light for their operation (Jacquinot, 1954), it has become a preferred instrument for use in ground-based optical remote sensing of the atmosphere (Hernandez and Killeen, 1986).

The FPI was first demonstrated by two French physicists, Charles Fabry and Alfred Perot, in 1897. The interferometer is a spectroscopic device of extremely high resolving power. The device, shown in Figure A.1, makes use of interference phenomena due to waves successively reflected between two plane, parallel, highly reflecting surfaces separated in space; these are usually glass optical plates with a multilayer dielectric film coating or an aluminium film coating.

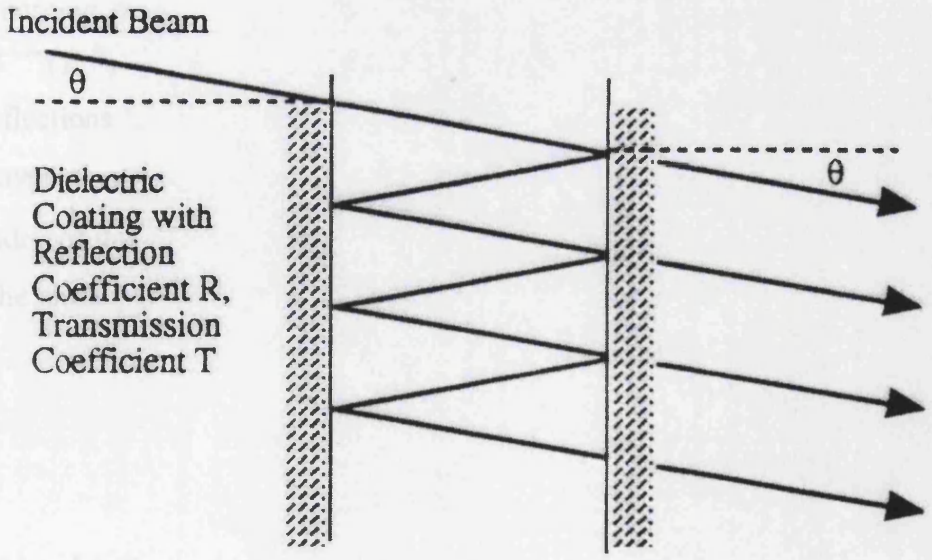


Figure A.1 Multiple reflections in a Fabry-Perot etalon (Lothian, 1975).

The enclosed air gap generally ranges from several millimeters to several centimeters. The uncoated sides of the plates are often made to produce a slight

wedge shape (~30'), as illustrated in Figure A.2, to reduce the interference pattern arising from reflections from these sides.

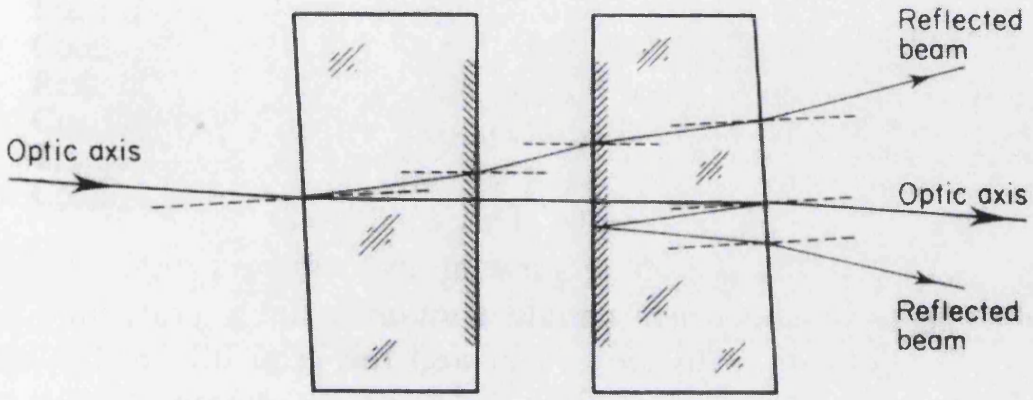


Figure A.2 Fabry-Perot Etalon with orientation of the plate wedge angles (Vaughan, 1989).

A.2 Theory of the Fabry - Perot Interferometer

Following the notation and treatment of Lothian (1975), if we consider the incidence of a light beam on the Fabry-Perot etalon, if the surface reflection and transmission coefficients are defined to be R and T respectively, then in terms of amplitudes, these coefficients are $R^{1/2}$ and $T^{1/2}$. From Figure A.3 it is seen that after emerging from the etalon the beam which has gone straight through has amplitude $(T^{1/2})(T^{1/2}) = T$ and the successive beams which have undergone 2, 4, 6 internal reflections have amplitudes TR, TR^2, TR^3 . Each of these successive beams will have travelled an extra distance $2\mu t\cos\theta$, where t is the plate separation, μ is the refractive index of the medium between the plates and θ is the angle of incidence on the plates. The phases of successive beams will differ by

$$\delta = \left(\frac{2\pi}{\lambda}\right) 2\mu t\cos\theta \qquad \text{(A.1)}$$

where λ is the wavelength of the plane wave incident on the etalon.

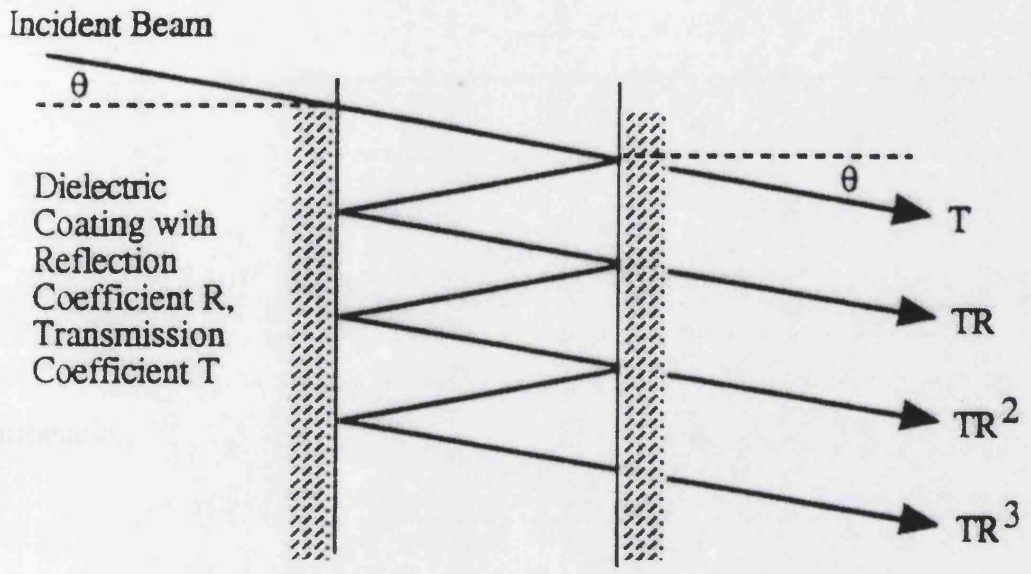


Figure A.3 Multiple reflections in a Fabry-Perot etalon, showing amplitudes of successive beams (Lothian, 1975).

If the incident wave is of the form

$$y = A\exp(i\omega t)$$

then the sum of the beams may be written

$$\begin{aligned} y &= T\exp(i\omega t) + TR\exp(i(\omega t - \delta)) + TR^2\exp(i(\omega t - 2\delta)) + \ldots \\ &= T\exp(i\omega t)\left[1 + R\exp(-i\delta) + R^2\exp(-2i\delta) + \ldots\right] \end{aligned} \tag{A.2}$$

where the terms ($\omega t - \delta$) and ($\omega t - 2\delta$) result from the change in phase experienced by each successive transmitted beam. This is an infinite geometrical progression and its sum is given as:

$$y = \left[\frac{T}{1 - R\exp(-i\delta)}\right]\exp(i\omega t) \tag{A.3}$$

The term in the square brackets is the amplitude, A. The intensity is AA^*, where A^* is the complex conjugate of A, and is therefore

$$
\begin{aligned}
I &= \frac{T^2}{[1-R\exp(-i\delta)][1-R\exp(i\delta)]} \\
&= \frac{T^2}{(1+R^2-R[\exp(-i\delta)+\exp(i\delta)])} \\
&= \frac{T^2}{1+R^2-2R\cos\delta}
\end{aligned}
\tag{A.4}
$$

Introducing the substitution

$$
\cos\delta = 1-2\sin^2\left(\frac{\delta}{2}\right) \tag{A.5}
$$

the following is obtained

$$
\begin{aligned}
I &= \frac{T^2}{(1-R)^2+4R\sin^2\left(\frac{\delta}{2}\right)} \\
&= \frac{T^2}{(1-R)^2}\left(\frac{1}{1+\frac{\left[4R\sin^2\left(\frac{\delta}{2}\right)\right]}{(1-R)^2}}\right)
\end{aligned}
\tag{A.6}
$$

which is known as the Airy function. Curves of I versus δ plotted from this equation are shown in Figure A.4, for different values of reflection coefficient, R . Since constructive interference can only occur when the path difference between successively transmitted beams is equal to an integral number of wavelengths, the maxima of intensity i.e. the peaks in Figure A.4 occur when

$$
2\mu t\cos\theta = n\lambda \tag{A.7}
$$

where n is an integer known as the order of interference. It can be seen that a wavelength $\lambda + \Delta\lambda$ satisfying the equation

$$2\mu t\cos\theta = (n-1)(\lambda + \Delta\lambda) \tag{A.8}$$

also produces fringes at the angular position θ. Multiplying out the RHS of Equation A.8

$$n\Delta\lambda = (\lambda + \Delta\lambda)$$

leads, using Equation A.7, to

$$\Delta\lambda = \frac{\lambda(\lambda + \Delta\lambda)}{2\mu t\cos\theta} \tag{A.9}$$

For near normal incidence, $\theta = 0$ and $\Delta\lambda << \lambda$ resulting in a simplified relation for

$$\Delta\lambda = \frac{\lambda^2}{2\mu t} \tag{A.10}$$

where $\Delta\lambda$ is called the Free Spectral Range of the interferometer and is the wavelength interval between successive peaks. The peaks will have a full width at half maximum (FWHM) given by

$$d\lambda_R = \Delta\lambda \frac{(1-R)}{\pi R^{\frac{1}{2}}} \tag{A.11}$$

which can be expressed as

$$d\lambda_R = \frac{\Delta\lambda}{N_R} \tag{A.12}$$

where

$$N_R = \frac{\pi R^{\frac{1}{2}}}{1-R} \tag{A.13}$$

N_R is known as the *reflecting finesse* and is dependent only on R. From the above equations, it is clear that if R increases, N_R increases, and the peaks become narrower as seen in Figure A.3.

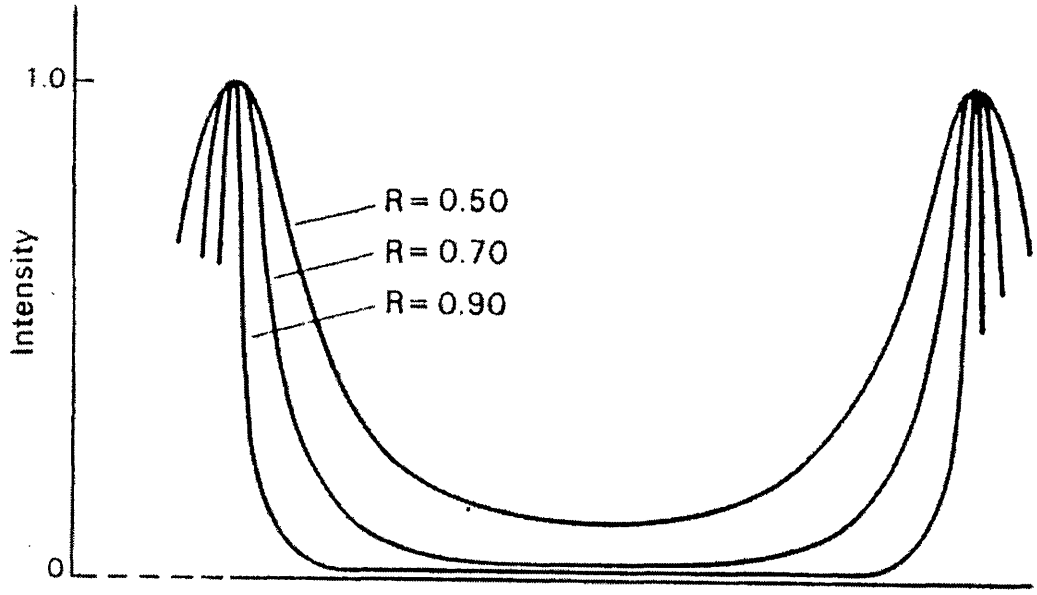


Figure A.4 Intensity distribution of Fabry-Perot fringes for three different values of reflection coefficient R (Lothian, 1975).

Examining the Airy function (Equation A.6) it is possible to find the peak transmission for the etalon to be

$$\tau_A = \left(\frac{T}{1-R} \right)^2 \tag{A.14}$$

So far it has been assumed $T + R = 1$. Being more realistic and taking absorption of some of the light by the reflective coating on the plates into account, $T = 1 - A - R$, where A is the absorptance of the plates. Rewriting the above equation gives

$$\tau_A = \left\{ 1 - \frac{A}{1-R} \right\}^2 \tag{A.15}$$

This indicates that while increasing the value of R will produce a high value of N_R, for a given value of A, it will also decrease the value of τ_A for the etalon. Thus the choice of R for the etalon is a compromise between high resolving power and high light throughput.

A.3 Non-ideal FPI behaviour

Defects encountered in practice with the FPI will include the plates not having perfectly flat surfaces or not being perfectly parallel and also that the instrument will be used over a finite angle of incidence range. All of these defects will produce broadening in the spectral line for the interferometer. To account for this, the instrument function includes a term known as the *defects function* (D), representing the imperfections in the etalon plates, and an *aperture function* (W), dealing with the effect of a finite range of incidence angles. The instrumental function is then given by the convolution of these terms with the Airy function (A)

$$E = A*D*W \tag{A.16}$$

The resolution of the interferometer is therefore limited to the width of the instrument function $d\lambda_E$. The effective resolving power is defined as

$$R_E = \lambda(d\lambda_E)^{-1} \tag{A.17}$$

and the ratio of the Free Spectral Range of the instrument to the instrumental line width is called the *Effective Finesse*, N_E, for the instrument

$$N_E = \Delta\lambda(d\lambda_E)^{-1} \tag{A.18}$$

An approximation for the Effective Finesse (Chabbal, 1953) is given by

$$\frac{1}{N_E^2} = \frac{1}{N_R^2} + \frac{1}{N_D^2} + \frac{1}{N_F^2} \tag{A.19}$$

where N_R, N_D and N_F are the reflection, defects and aperture finesses respectively. N_D is often called the limiting finesse as both N_R and N_F can be made to have relatively high values, without compromising light throughput. This is because it is not possible to increase N_D above the value $x/2$ if the plates are flat to λ/x.

A.4 Low light levels and the FPI

An important consideration in the use of an FPI in spectroscopic mode with a faint source is the luminosity (L) of the device, which is defined as the ratio of the flux falling on the instrument to the luminance of the source (Jacquinot, 1954). For an FPI the luminosity can be defined as

$$L = S\Omega\tau_E \tag{A.20}$$

where S is the surface area of the plates, Ω is the solid angle of acceptance and τ_E is the transmission factor for the instrument which is defined as

$$\tau_E = \tau_A\tau_D\tau_F\tau_O \tag{A.21}$$

where τ_A is the transmission factor due to the coatings, τ_D is the transmission factor due to the convolution of the Airy and defects functions and τ_F is the transmission factor due to the convolution of the convolved Airy and defects functions with the aperture function and τ_O is the transmission factor due to all other optics for the FPI including filters and lenses.

It has been shown that while the finesses N_R, N_D and N_F can be combined in many ways to achieve a certain resolution, only one combination will give the maximum luminosity (Jacquinot, 1954). In general the FPIs designed and built by the Atmospheric Physics Laboratory at UCL have been intended for use in observing the 630nm emission from atomic oxygen in the thermosphere, however the more intense atomic oxygen line at 557.7nm, originating from the lower thermosphere has also been observed using these instruments as has the 843nm hydroxyl doublet. These nightglow emission features have intensities which may differ by more than an order of magnitude which leads to the need for a compromise between resolution at the higher intensities and luminosity at the lower intensities. These factors need to be considered when choosing the reflectivity of the etalon coatings and also the spacing of the etalon gap.

A.5 Comparative advantage of the FPI

Comparison of the FPI with other spectroscopic devices was carried out by Jacquinot and Dufour (1948) and Jacquinot (1954), and examination of their findings explains the reasons for its use in remote sensing of the weak terrestrial airglow. Vaughan (1989) summarises these findings by representing the flux transmitted, $\Im$, by three types of spectrometer (prism, grating and plane Fabry-Perot) for a quasi-monochromatic line (narrow compared with the resolving limit) in the following manner

$$\Im = \frac{1}{\Re} BTKSC \tag{A.22}$$

where $\Re$ is the effective resolving power of the spectrometer, S is the area of the dispersive element (prism base, effective grating surface or usable Fabry-Perot plate area), K is the luminosity factor (which Vaughan shows to be quite close to unity) and C has a different value for each type of spectrometer. B is the spectral radiance of the source and T is the peak transmission of the spectral filter at its centre.

The flux gathering performance of the different spectrometers may be compared from the respective values of C. These were conveniently summarized by Chabbal and Jacquinot (1955) as follows:

(i) for the prism $C = \lambda(dn/d\lambda)\beta \approx 1/500$ where β is the angular height of the spectrometer slit;

(ii) for the grating $C = \beta \approx 1/50$; and

(iii) for the plane Fabry-Perot spectrometer $C = 2\pi$

This shows that for instruments of similar resolving power and comparable size (effective area S), the grating instrument is an order of magnitude more efficient than the prism spectrometer and more than two orders of magnitude less efficient than the Fabry-Perot spectrometer.

For this reason the FPI is the instrument of choice for applications of high resolution, for example when the source used is weak, as is the case in airglow remote sensing. The Michelson interferometer has been shown (Thorne, 1988) to have similar luminosity characteristics to the FPI, but this interferometer is usually used in applications of lower resolution with a broader range of wavelengths.

Appendix B

The following is the Matlab script used to derive values of *hmF2* from EISCAT CP experiments stored in NCAR format at RAL.

```
%_____________________/////////////////////////////////////
%
%  Matlab Function: hoighty.m
%
%  Usage: [ne,alt,ut,az,el] = hoighty(EISCAT_FILENAME);
%
%/////////////////_____________________________________________


function [ne,alt,ut,az,el] = hoighty(EISCAT_FILENAME)

   % Use a modified version of ncar_sort2 to return the
   % AZ and EL along with the usual parameters
   [NE, ALT, SNR, TE, TI, VEL, AZ, EL, prol, time,...
    date] = ncar_sort2(EISCAT_FILENAME);


   % Get the peak density
   [val, NEmax] = max(NE);

   % Remove all NaNs and values greater than 10e13
   % from the data set
   profiles = find(~isnan(val) & val < 10e13);
   NEmax = NEmax(profiles);

   order = 4
   model_hts = [100000:100:500000];
   for i=1:length(NE)
     p = polyfit(ALT(:,i), NE(:,i),order);
     NE_values = polyval(p,model_hts);
     [peak_NE(i),pos] = max(NE_values);
     peak_ALT(i) = model_hts(pos);
   end

   % Reduce the data from the above criteria
   tempNE = peak_NE(profiles);
   tempALT = peak_ALT(profiles);
   tempTime = time(profiles,:);
   az = AZ(profiles);
   el = EL(profiles);

   % Make the UT time continuous for the whole experiment run
   ut = ncar_hours(tempTime);

   % copy the peak density and its corresponding altitude
   ne = peak_NE;
   alt = peak_ALT;

end;
```

Appendix C

The following is the Matlab script used to calculate the correction necessary for the influence of electric fields on neutral winds derived from values of *hmF2* from EISCAT CP experiments.

```
function [outimes,outwinds] = doit(times,etimes,winds,velcor)

%  This script takes EISCAT ion velocity data and derived neutral
%  winds from the same EISCAT experiments to produce derived neutral
%  winds with electric field corrections.

 elcor = velcor/0.9763;

 x = length(times);
 y = length(etimes);

 for i = 1:x
      outwind(i) = 0;
      for j = 1:y
            if (abs(etimes(j) -times(i)) < .2);
                  outwind(i) = winds(i) + elcor(j);
            end;
      end;
      if ((outwind(i) == 0) | (isnan(outwind(i))));
            outwind(i) = winds(i);
      end;
 end;

 outimes = times;
 outwinds = reshape(outwind,length(outwind),1);

end;
```

Akasofu, S. I. (1987) in *The solar wind and the Earth* Ed. Akasofu, S. I. and Y. Kamide, Terra Scientific Publishing Company, Tokyo.

Anderson, D. N., M. J. Buonsanto, M. Codrescu, D. Decker, C. G. Fesen, T. J. Fuller-Rowell, B. W. Reinisch, P. G. Richards, R. G. Roble, R. W. Schunk and J. J. Sojka (1998), Intercomparison of physical models and observations of the ionosphere. *J. Geophys. Res.*, **103**, 2179-2192.

Appleton, E. V. and M. A. F.Barnet (1925) *Proc. Roy. Soc. A*, **109**, 621-41.

Armstrong, E. B. (1956) The observation of line profiles in the airglow and aurora with a photoelectric Fabry-Perot interferometer in *The airglow and the aurorae* Ed. E. B. Armstrong and A. Dalgarno, 366-373, Pergamon Press, Oxford.

Armstrong, E. B. (1958) Largeur de la raie OI λ 5577 du ciel nocturne et des aurores. *J. Phys. Radium*, **19**, 358-365.

Aruliah, A.L. and D. Rees (1995), The trouble with thermospheric vertical winds : geomagnetic, seasonal and solar cycle dependence at high latitudes. *J. Atmos. Terr. Phys.*, **57**, 597-609.

Aruliah, A.L., A. D. Farmer, D. Rees and U. Brandstrom (1996a), The seasonal behaviour of the high-latitude thermospheric winds and ion velocities observed over one solar cycle. *J. Geophys. Res.*, **101**, 15,701-15,711.

Aruliah, A. L., A. D. Farmer, T. J. Fuller-Rowell, M. N. Wild, M. Hapgood and D.Rees (1996b) An equinoctial asymmetry in the high-latitude thermosphere and ionosphere. *J. Geophys. Res.*, **101**, 15,713-15,722.

Aruliah, A.L., J. Schoendorf, A. D. Aylward and M.N. Wild (1997), Modeling the high-latitude equinoctial asymmetry. *J. Geophys. Res.*, **102**, 27,207-27,216.

Aruliah, A. L., I. C. F. Mueller-Wodarg and J. Schoendorf (1999) Consequences of geomagnetic history on the high-latitude thermosphere and ionosphere: Averages, *J. Geophys. Res.*, **104**, 28073-28088, 1999

Babcock, H. D. (1923), A study of the green auroral line by the interference method. *Astrophys. J.*, **57**, 209-221.

Balan, N., Y. Otsuka, G.J. Bailey and S. Fukao (1998), Equinoctial asymmetries in the ionosphere and thermosphere observed by the MU radar. *J. Geophys. Res.*, **103**, 9481-9495.

Banks, P. (1966), Collision frequencies and energy transfer: ions. *Planet. Space Sci.*, **14**, 1105-1122

Banks P. M. and G. Kockarts (1973) *"Aeronomy"*, 2 volumes (Parts A and B). Academic Press.

Bibl, K. and B. W. Reinisch (1978), The universal digital ionosonde. *Radio Sci.*, 13, 519-530.

Biondi, M. A., S. Y. Sazykin, B. G. Fejer, J. W. Meriwether and C. G. Fesen (1999), Equatorial and low latitude thermospheric winds: Measured quiet time variations with season and solar flux from 1980 to 1990. *J. Geophys. Res.*, **104**, 17,091-17,106.

Bradley, P. A. and J. R. Dudeney (1973), A simple model of the vertical distribution of electron concentration in the ionosphere. *J. Atmos. Terr. Phys.*, **35**, 2131-2146.

Brasseur G. and S. Solomon (1986) *Aeronomy of the Middle Atmosphere* – 2nd Edition. D. Reidel Publishing Company, Dordrecht, Holland.

Breit G. and M. Tuve (1926) *Phys. Rev.*, **28**, 554-573.

Buisson, H., C. Fabry and H. Bourget (1914) An application of interference to the study of the Orion nebula. *Astrophys. J.*, **40**, 241-258.

Buonsanto, M. J., J. E. Salah, K. L. Miller, W. L. Oliver, R. G. Burnside and P. G. Richards (1989), Observations of neutral circulation at mid-latitudes during the Equinox Transition Study. *J. Geophys. Res.*, **94**, 16,987-16,997.

Buonsanto, M. J. (1990) Observed and calculated F2 peak heights and derived meridional winds at mid-latitudes over a full solar cycle. *J. Atmos. Terr. Phys.*, **52**, 223-240.

Buonsanto, M. J. (1991) Neutral winds in the thermosphere at mid-latitudes over a full solar cycle: A tidal decomposition. *J. Geophys. Res.*, **96**, 3711-3724.

Buonsanto, M. J., Y. K. Tung and D. P. Sipler (1992), Neutral atomic oxygen density from nighttime radar and optical wind measurements at Millstone Hill. *J. Geophys. Res.*, **97**, 8673-8680.

Buonsanto, M. J., M. Codrescu, B. A. Emery, C. G. Fesen, T. J. Fuller-Rowell, D. J. Melendez-Alvira and D. P. Sipler (1997a) Comparison of models and measurements at Millstone Hill during the January 24-26, 1993, minor storm interval. *J. Geophys. Res.*, **102**, 7267-7277.

Buonsanto, M. J., M. J. Starks, J. E. Titheridge, P. G. Richards and K. L. Miller (1997b), Comparison of techniques for derivation of neutral meridional winds from ionospheric data. *J. Geophys. Res.*, **102**, 14,477-14,484.

Buonsanto, M. J., D. P. Sipler, G. B. Davenport and J. M. Holt (1997c), Estimation of the O^+,O collision frequency from coincident radar and Fabry-Perot observations at Millstone Hill. *J. Geophys. Res.*, **102**, 17,267-17,274.

Buonsanto, M. J. and O. G. Witasse (1999) An updated climatology of thermospheric neutral winds and F region ion drifts above Millstone Hill. *J. Geophys. Res.*, **104**, 24,675-24,687.

Burns, A. G., T. L. Killeen and R. G. Roble (1991) A theoretical study of thermospheric composition perturbations during an impulsive geomagnetic storm. *J. Geophys. Res.*, **96**, (A8), 14,153-14,167.

Burnside, R. G., C. A. Tepley and V. B. Wickwar (1987), The O+-O collision cross-section: can it be inferred from aeronomical measurements? *Annales Geophys.*, **5A**, 343-350.

Caplan, J. (1975) Temperature and pressure effects on pressure scanned etalons and gratings. *Appl. Opt.*, **14**, 1585-1591.

Cabannes J. and J. Dufay (1956) *Rev. d'Opt.*, **35**, 103.

Chabbal, R.(1953) Recherche des meillures conditions d'utilisation d'un spectrometre photoelectrique Fabry-Perot. *J. Rech. CNRS*, **24**, 138-186.

Chabbal, R. and P. Jacquinot (1955) L'utilisation de l'interferometre Fabry-Perot. *Nuovo Cimento Suppl.*, **2** Ser X, 661-669.

Chanin, M. L. and A. Hauchecorne (1991), Lidar study of the structure and dynamics of the middle atmosphere. *Indian J. Radio and Space Phys.*, **20**, 1-11.

Chapman, S. (1950) Upper atmosphere nomenclature. *Bull. Amer. Met. Soc.*, **31**, 288-290

Chiu, Y. T. (1975), An improved phenomenological model of ionospheric density. *J. Atmos. Terr. Phys.*, **37**, 1563-1570.

Codrescu, M. V., T. J. Fuller-Rowell, J. C. Foster, J. M. Holt and S. J. Cariglia (2000) Electric field variability associated with the Millstone Hill electric field model. *J. Geophys. Res.*, **105**, 5265-5273.

Colerico, M., M. Mendillo, D. Nottingham, J. Baumgardner, J. Meriwether, J. Mirick, B. W. Reinisch, J. L. Scali, C. G. Fesen and M. A. Biondi (1996) Coordinated measurements of F region dynamics related to the thermospheric midnight temperature maximum. *J. Geophys. Res.*, **101**, 26783-26793.

Conde, M. and R. W. Smith (1997) "Phase compensation" of a separation scanned, all-sky imaging Fabry-Perot spectrometer for auroral studies. *Appl. Opt.*, **36**, 5441-5450.

COSPAR International Reference Atmosphere (1972) , Solar Geophysical Data , published
World Data Center A, Boulder, Colorado.

Creekmore S. P., J. M. Straus, R. M. Harris, B. K. Ching and Y. T. Chiu, (1975) A global model of thermospheric dynamics. I. Wind and density field derived from phenomenological temperature. *J. Atmos Terr. Phys.* , **36**, 491-515.

Dalgarno, A. (1964), Ambipolar diffusion in the F-region. *J. Atmos. Terr. Phys.*, **26**, 989.

Davis, C. J. (1993), *The interaction between the thermosphere and ionosphere at high latitudes*, Ph.D. Thesis, University of Southampton, U.K.

Davis C. J., A. D. Farmer and A. Aruliah (1995) An optimised method for calculating the O^+-O collision parameter from aeronomical measurements. *Ann. Geophysicae* , **16**, 1138-1143.

Davis, C. J. (1999) Private Communication.

Dickinson R. E. et al. (1975) Meridional circulation in the thermosphere I, Equinox conditions. *J. Atmos. Sci.*, **32**, 1737-1754.

Dickinson, R. E., E. C. Ridley and R. G. Roble (1981), A three dimensional general circulation model of the thermosphere. *J. Geophys. Res.*, **86**, 1499-1512.

Doyle, D. B. (1987) Interferometric measurements of high latitude thermospheric winds from ground and space. *M. Sc. Thesis*, St. Patricks College, Maynooth, Ireland.

Dougherty, J. P. and D. T. Farley (1960), A Theory of Incoherent Scattering of Radio Waves by a Plasma. *Proc. Roy. Soc.*, **A_259**, 79-99.

Duboin, M.-L., M. Lafeuille, A. Richmond and P. Rouchette (1988) Etude statistique des vents neutres en region F au-dessus de Saint-Santin. *Doc. De Trav. CNET/CRPE/1158*, Cent. Natl. d'Etud. Des Telecommun., Issy-les-Moulineaux, France.

Duboin, M.-L. and M. Lafueille (1992) Thermospheric dynamics above Saint-Santin: statistical study of the data set. *J. Geophys. Res.*, **97**, 8661-8671.

Dudeney, J. R. (1983), The accuracy of simple methods for determining the height of the maximum electron concentration of the F2-layer from scaled ionospheric characteristics. *J. Atmos. Terr. Phys.*, **45**, 629-640.

Duncan, R. A. (1956). *Aust. J. Phys.*, **9**, 436.

Dyson, P. L., T. P. Davies, M. L. Parkinson, A. J. Reeves, P. G. Richards, C. E. Fairchild (1997), Thermospheric neutral winds at southern mid-latitudes: A comparison of optical and ionosonde *hmF2* methods. *J. Geophys. Res.*, **102**, 27189-27196.

Emery, B. A. (1978) Neutral thermospheric winds deduced above Millstone Hill, 2, Seasonal wind variations, 1970-1971. *J. Geophys. Res.*, **83**, 5704-5716.

Fabry, C. and H. Buisson (1914) Verification experimentale du principe de Doppler-Fizeau. *Compt. Rend.*, **158**, 1498-1499.

Farmer, A. D. (1987), Extended abstract from the MIST meeting Edinburgh, April 1986. *Q. J. Roy Astron. Soc.*, **28**, 43-53.

Farmer, A. D., K. J. Winser, A. Aruliah and D. Rees (1990), Ion-neutral dynamics: comparing Fabry-Perot measurements of neutral winds with those derived from radar observations. *Adv. Space Res.* , **10**, 281.

Fauliot, V., G. Thuiller and M. Herse (1993) Observation of the F-region horizontal and vertical winds in the auroral zone. *Annales Geophys.*, **11**, 17-28.

Fejer, J. (1960). *Can. J. Phys.*, **38**, 1114-1133.

Fejer, B. G. and L. Schierliess (1997) Empirical models of storm time equatorial zonal electric fields. *J. Geophys. Res.*, **102**, 24047-24056.

Field, P. R. and H. Rishbeth (1997), The response of the ionospheric F2-layer to geomagnetic activity: an analysis of worldwide data. *J. Atmos. Solar-Terr. Phys.*, **59**, 163-180.

Foster, J. C. (1984), Ionospheric signatures of magnetospheric convection. *J. Geophys. Res.*, **89**, 855-865.

Fuller-Rowell, T. J. and D. Rees (1980) A three-dimensional, time dependent global model of the thermosphere. *J. Atmos. Sci.*, **37**, 2545-2657.

Fuller-Rowell, T. J. and D. Rees (1981) A three-dimensional, time dependent simulation of the thermosphere to a geomagnetic substorm. *J. Atmos. Terr. Phys.*, **43**, 701-721.

Fuller-Rowell, T. J., S. Quegan, D. Rees, R. J. Moffett, G. J. Bailey (1984), The effect of realistic conductivities on the high-latitude neutral thermospheric circulation. *Planet. Space Sci.*, **32**, 469-480.

Fuller-Rowell, T. J., D. Rees, S. Quegan, R. J. Moffett, G. J. Bailey (1988), Simulations of the seasonal and UT variations of the thermosphere and ionosphere using a coupled, three-dimensional, global model. *Pure Appl. Geophys.*, **127**, 189-217.

Gagne, J. M., J. P. Saint-Dizier and M. Picard (1974) Methode d'echantillonnage des fonctions deterministes en spectroscopie: applicaton a un spectrometre multicanal par comptage photonique. *Appl. Opt.*, **13**, 581-588.

Gordon, W. E. (1958). *Proc. I. R. E.*, **46**, 1824.

Greenwald, R. A., K. B. Baker, R. A. Hutchins and C. Hanuise (1985) An HF phased-array radar for studying small scale structure in the high latitude ionosphere. *Radio Sci.*, **20**, 63-79.

Gulledge, I. S., D. M. Packer, S. G. Tilford and J. T. Vanderslice (1968) *J. Geophys. Res.*, **73**, 5535.

Gurubaran, S. and R. Sridharan (1993), Effect of meridional winds and neutral temperatures on the F layer heights over low latitudes. *J. Geophys. Res.*, **98**, (A7), 11629-11635.

Hagan, M. E. (1993) Quiet time upper thermospheric winds over Millstone Hill between 1984 and 1990. *J. Geophys. Res.*, **98**, 3731-3739.

Hagfors, T. (1961). *J. Geophys. Res.*, **66**, 1699-1712.

Hargreaves, J. K. (1979) *The Upper Atmosphere and Solar Terrestrial Relations*. Van Nostrand Reinhold, New York.

Hargreaves, J. K. (1992) *The Solar Terrestrial Environment*. Cambridge University Press.

Hays, P. B., T. L. Killeen, N. W. Spencer, L. E. Wharton, R. G. Roble, B. A. Emery, T. J. Fuller-Rowell, D. Rees, L. A. Frank, and J. D. Craven (1984), Observations of the dynamics of the polar thermosphere. *J. Geophys. Res.*, **89**, 5597-5612.

Hedin, A. E. (1987) MSIS-86 Thermospheric model. *J. Geophys. Res.*, **92**, (A5), 4649-4662.

Hedin, A. E. (1988) Atomic oxygen modelling in the upper thermosphere. *Planet. Space Sci.*, **36**, 907-920.

Hedin, A. E., N. W. Spencer and T. K. Killeen (1988), Empirical global model of upper thermosphere winds based on Atmosphere and Dynamics Explorer satellite data. *J. Geophys. Res.*, **93**, 9959-9978.

Hedin, A. E., M. A. Biondi, R. G. Burnside, G. Hernandez, R. M. Johnson, T. L. Killeen, C. Mazaudier, J. W. Meriwether, J. E. Salah, R. J. Sica, R. W. Smith, N. W. Spencer, V. B. Wickwar and T. S. Virdi (1991), Revised global model of thermosphere winds using satellite and ground-based observations. *J. Geophys. Res.*, **96**, 7657-7688.

Hedin, A. E., M. J. Buonsanto, M. Codrescu, M.-L. Duboin, C. G. Fesen, M. E. Hagan, K. L. Miller and D. P. Sipler (1994) Solar activity variations in midlatitude thermospheric meridional winds. *J. Geophys. Res.*, **99**, 17,601-17,608.

Hedin, A. E., E. L. Fleming, A. H. Manson, F. J. Schmidlin, S. K. Avery, R. R. Clark, S. J. Franke, G. J. Fraser, T. Tsuda, F. Vial and R. A. Vincent (1996) Empirical wind model for the upper, middle and lower atmosphere. *J. Atmos. Terr. Phys.*, **58**, 1391-1403

Heppner, J. P. (1977) Empirical models of high latitude electric field. *J. Geophys. Res.*, **82**, 115-1125.

Hernandez, G. and R.G. Roble (1976) Direct measurements of nighttime thermospheric winds and temperatures, 1, seasonal variations during geomagnetic quiet periods. *J. Geophys. Res.*, **81**, 2065.

Hernandez, G. (1986) *Fabry-Perot Interferometers*. Cambridge University Press.

Hernandez, G. and T. L. Killeen (1986) Optical measurements of winds and temperatures in the upper atmosphere. *CIRA-86*, COSPAR International Reference Atmosphere.

Houghton, J. T. (1977) *The Physics of Atmospheres.* Cambridge University Press.

IAGA Working Group (1996) *IGRF 1995: revision*, EOS, AGU **77**, 153..

Igi, S., W. L. Oliver and T. Ogawa (1999) Solar cycle variations of the thermospheric meridional wind over Japan derived from measurements of h_mF_2 . *J. Geophys. Res.*, **104**, 22427-22431.

Jacquinot, P. and C. Dufour (1948) Conditions optiques d'emploi des cellules photo-electriques dans les spectrographes et les interferometres. *J. Res. CRNS*, **6**, 91-103.

Jacquinot, P. (1954) The luminosity of spectrometers with prisms, gratings or Fabry-Perot etalons. *J. Opt. Soc. Am.*, **44**, 761-765.

Jarvis, M. J., B. Jenkins and G. A. Rodgers (1998) Southern hemisphere observations of a long-term decrease in F region altitude and thermospheric wind providing possible evidence for global thermospheric cooling. *J. Geophys. Res.*, **103**, 20774-20787.

Khachikjan, G. Ja., A. I. Pogoreltsev and Ja. V. Drobjeva (1997) Mean thermospheric winds at middle latitudes of the northern hemisphere derived from longitudinal variations of the F2-layer peak height. *J. Atmos. Solar-Terr. Phys.*, **59**, 1391-1403.

Killeen, T. L., P. B. Hays, B. C. Kennedy and D. Rees (1982) Stable and rugged etalon for the Dynamics Explorer Fabry-Perot interferometer 2: Performance. *Appl. Opt.* , **21**, 3903-3912.

Killeen, T. L., P. B. Hays, N. W. Spencer and L. E. Wharton (1983), Neutral winds in the polar thermosphere from Dynamics Explorer. *Adv. Space Res.*, **2**, 133-136.

Killeen, T. L. and R. G. Roble (1984) An analysis of the high-latitude thermospheric wind pattern calculated by a Thermospheric General Circulation Model 1. Momentum Forcing. *J. Geophys. Res.*, **89**, 7509-7522.

Killeen, T. L., R. G. Roble and N. W. Spencer (1987), A computer model of global thermospheric winds and temperatures. *Adv. Space Res.*, **7**, 207.

Killeen, T. L. and R. G. Roble (1988) Thermosphere dynamics: Contributions from the first 5 years of the Dynamics Explorer program. *Rev. of Geophys.*, **26**, 329-367.

Killeen, T. L. (1990) *Upper Atmosphere Physics Coursebook.* Space Physics Research Laboratory, University of Michigan, Ann Arbor.

Killeen, T. L., Y.-I. Won, R. J. Niciejewski and A. G. Burns (1995) Upper thermosphere winds and temperatures in the geomagnetic polar cap: Solar cycle, geomagnetic activity, and interplanetary magnetic field dependencies. *J. Geophys. Res.*, **100**, 21327-21342.

Kohl H. and J.W. King (1967) Atmospehric winds between 100 and 700 km and their effect on the ionosphere. *J. Atmos. Terr. Phys.*, **29**, 1045-1062.

Kosch, M. J., T. Hagfors and D. Rees (1997) A new Fabry-Perot interferometer for atmospheric studies with the EISCAT incoherent radar. *Adv. Space Res.*, **20**, 1133-1136.

Lal, C. (1996) Seasonal trend of geomagnetic activity derived from solar-terrestrial geometry confirms an axial-equinoctial theory and reveals deficiency in planetary indices. *J. Atmos. Terr. Phys.*, **58**, 1497-1506.

Lathuillere, C., J. Lilensten, W. Gault and G. Thuillier (1997) Meridional wind in the auroral thermosphere: results from EISCAT and WINDII-O(1D) coordinated measurements. *J. Geophys. Res.*, **102**, 4487-4492.

Lee, J. S., J. P. Doering, T. A. Potemra and L. H. Brace (1980), Measurements of the ambient photoelectron spectrum from Atmosphere Explorer, I, AE-E measurements below 300 km during solar minimum conditions. *Planet. Space Sci.*, **28**, 947.

Lehtinen M. and A.-L. Turunen (1981) EISCAT UHF antenna calibration. *EISCAT Technical Note 81/30.*

Lilensten, J. and C. Lathuillere (1995) The meridional thermospheric neutral wind measured by the EISCAT radar. *J. Geomagn. Geoelectr.* , **47**, 911-920.

Lothian, G. F. (1975) *Optics and its uses*, Van Nostrand Reinhold.

McEwan, M. J. and L. F. Philips (1975) *Chemistry of the Upper Atmosphere*. Edward Arnold Ltd., London.

McWhirter, I., D. Rees and A. H. Greenaway (1982), Miniature Imaging photon detetctors III. An assessment of the performance of the resistive anode IPD. *J. Phys. E.: Sci. Instrum.*, **15**, 145-150.

Mayr, H. G. and H. Volland, (1972) Theoretical model for the latitude dependence of the thermospheric annual and semi-annual variations. *J. Geophys. Res.*, **77**, 6774-6790.

Meriwether, J. W., T. L. Killeen, F. G. McCormac, A. G. Burns and R. G. Roble (1988) Thermospheric winds in the geomagnetic polar cap for solar minimum conditions *J. Geophys. Res.*, **93**, 7478-7492.

Miller, K. L., D. G. Torr and P. G. Richards (1986), Meridional winds in the thermosphere derived from measurements of F2 layer height. *J. Geophys. Res.*, **91**, 4531-4535.

Miller, K. L., J. E. Salah and D. G. Torr (1987), The effects of electric fields on measurements of meridional neutral winds in the thermosphere. *Annales Geophys.*, **5A**, 337-342.

Miller, K. L., R. L. Breninger, P. G. Richards and D. G. Torr (1989) The solar cycle variation of meridional neutral winds in the thermosphere (abstract). *Eos Trans. AGU*, **70**, 1243.

Miller, K. L., A. E. Hedin, P. J. Wilkinson, D. G. Torr and P. L. Richards (1990) Neutral winds derived from IRI parameters and from the HWM87 wind model for the Sundial Campaign of September, 1986. *Adv. Space Res.*, **10**, 99-102

Miller, K. L., P. G. Richards and H. Y. Wu (1993), A global-scale study of meridional winds and electron densities in the F-region during the SUNDIAL 1987 campaign. *Annales Geophys.*, **11**, 572-584.

Miller, K. L., M. Lemon and P. G. Richards (1997), A meridional wind climatology from a fast model for the derivation of meridional winds from the height of the ionospheric F2 region. *J. Atmos. Solar-Terr. Phys.*, **59**, 1805-1822.

Moffet, R. J., R. Sellek and G. J. Bailey (1990), The influence of the O+-O collision frequency on ionospheric F-region behaviour. *J. Atmos. Terr. Phys.*, **52**, 125-132.

Muller-Wodarg, I. C. F., A. D. Aylward and T. J. Fuller-Rowell (1999) Tidal oscillations in the thermosphere: A theoretical investigation of their sources. *Submitted for publication in J. Atmos. Solar-Terr. Phys.special IUGG issue.*

Nagy, A. F. and P. M. Banks (1970), Photoelectron fluxes in the ionosphere. *J. Geophys. Res.*, **75**, 6260.

Namgaladze, A. A., A. N. Namgaladze and M. A. Volkov (1996) Numerical modelling of the thermospheric and ionospheric effects of magnetospheric processes in the cusp region. *Ann. Geophysicae* , **14**, 1343-1355.

Narayaran, R., J. N. Desai, N. K. Modi, R. Raghavarao and R. Sridharan (1989) Dayglow photmetry: a new approach. *Appl. Opt.*, 28, 2138-2142.

Niciejewski, R., T. L. Killeen and M. Turnbull (1994) Ground-based Fabry-Perot interferometry of the terrestrial nightglow with a bare charge-coupled device: Remote field site deployment. *Opt. Eng.*, **33**, 457.

Omidvar, K., R. Menard and M. J. Buonsanto (1998) Empirical determination of the O^+O collision frequency. *J. Atmos. Solar-Terr. Phys.*, **60**, 1485-1496.

Parrat, L. G. (1961) *Probability and Experimental Errors in Science*, Chapter 5, John Wiley and Sons, New York..

Peterson, R. N., L. L. cogger, H. T. Meredith and G. G. Shepherd (1979) A laboratory discharge lamp source for the O_2 Atmospheric bands and OI 5577 A and 6300 A lines of a known line width. *Planet. Space Sci.*, **27**, 1209-1212.

Quegan, S., G. J. Bailey, R. J. Moffett, R. A. Heelis, T. J. Fuller-Rowell, D. Rees, R. W. Spiro (1982), Theoretical study of the distribution of ionization in the high-latitude ionosphere and the plasmasphere: First results on the mid-latitude trough and the light-ion trough. *J. Atmos. Terr. Phys.*, **44**, 619-640.

Ratcliffe, J. A. (1972), *An Introduction to the Ionosphere and Magnetosphere.* Cambridge University Press.

Rees, D., T. J. Fuller-Rowell and R. W. Smith (1980), Measurements of high-latitude thermospheric winds by rocket and ground-based techniques and their interpretation using a three-dimensional, time-dependent, thermospheric model. *Planet. Space Sci.*, **28**, 919-932.

Rees, D., I. McWhirter, P. B. Hays and T. Dines (1981) A stable, rugged, capacitance-stabilized piezoelectric scanned Fabry-Perot etalon. *J. Phys. E. : Sci. Instrum.*, **14**, 1320-1325.

Rees, D., N. Lloyd, P. J. Charleton, M. Carlson, J. Murdin and I. Haggstrom (1984), Comparison of plasma flow and thermospheric circulation over Northern Scandinavia using EISCAT and a Fabry-Perot interferometer. *J. Atmos. Terr. Phys.*, **46**, 545-564.

Richards, P. G. and D. G. Torr (1988), Ratios of photoelectron to EUV ionization rates for aeronomic studies. *J. Geophys. Res.*, **93**, 4060-4066.

Richards, P. G. (1991), An improved algorithm for determining neutral winds from the height of the F2 peak electron density. *J. Geophys. Res.*, **96**, 17839-17846.

Richmond, A. D. and S. Matsushita, (1975) Thermospheric response to a magnetic substorm. *J. Geophys. Res.*, **80,** 2839-2850.

Richmond, A. D., (1979), Large amplitude gravity wave energy production and dissipation in the thermosphere. *J. Geophys. Res.*, **84,** 1880-1890.

Richmond, A. D., M. Blanc, B. A. Emery, R. H. Wand, B. G. Fejer, R. F. Woodman, S. Ganguly, P. Amayenc, R. A. Behnke, C. Calderon and J. V. Evans (1980) An empirical model of quiet-day ionospheric electric fields at middle and low latitudes. *J. Geophys. Res.*, **85**, 4658-4664.

Richmond, A. D., E. C. Ridley and R. G. Roble (1992), A thermosphere/ionosphere general circulation model with coupled electrodynamics. *Geophys. Res. Lett.*, **19**, 601-604.

Rishbeth, H. and C. S. G. K. Setty (1961), The F-layer at sunrise. *J. Atmos. Terr. Phys.*, **20**, 263-276.

Rishbeth, H. (1967), The effects of winds on the ionospheric F2-peak. *J. Atmos. Terr. Phys.*, **29**, 225-238.

Rishbeth, H. and O. Garriot (1969) Introduction to Ionospheric Physics. Academic Press.

Rishbeth, H., S. Ganguly and J. C. G. Walker (1978) Field-aligned and field-perpendicular velocities in the ionospheric F2-layer. *J. Atmos. Terr. Phys.*, **40**, 767-784.

Rishbeth, H. and P. J. S. Williams (1985), The EISCAT ionospheric radar: The system and its early results. *Q. Jl R. Astr. Soc.*, **26**, 478-512.

Roach, F. E. and J. L. Gordon (1973) *The light of the night sky*. D. Reidel Publishing Company, Dordrecht, Holland.

Roble, R. G., R. E. Dickinson and E. C. Ridley (1982), The global circulation and temperature structure of the thermosphere with high-latitude plasma convection. *J. Geophys. Res.*, **87**, 1599-1614.

Roble, R. G. (1988), Impact of uncertainties in the O+-O collisional cross-section on the global structure and dynamics of the thermosphere and ionosphere. *EOS Transactions, AGU*, **69**, 417.

Roble, R. G., T. L. Killeen, N. W. Spencer, R. A. Heelis, P. H. Reiff and J. D. Winningham (1988a) Thermospheric dynamics during November 21-22, 1981: Dynamics Explorer measurements and thermospheric general circulation model predictions. *J. Geophys. Res.*, **93**, 209-225.

Roble, R. G., E. C. Ridley, A. D. Richmond and R. E. Dickinson (1988b), A coupled thermosphere/ionosphere general circulation model. *Geophys. Res. Lett.*, **15**, 1325-1328.

Roble, R. G., M. Kelley and C. C. Gardner (1992) Global change: Upper atmospheric research and the role of the NSF CEDAR program. *The Cedar Post*, No. 15, 1-4.

Russell, C. T. (1987) The magnetosphere, in *The solar wind and the Earth* Ed. Akasofu, S. I. and Y. Kamide, Terra Scientific Publishing Company, Tokyo.

Salah, J. E. and J. M. Holt (1974), Midlatitude thermospheric winds from incoherent scatter radar and theory. *Radio Sci.*, **9**, 301-313.

Salah, J. E. (1993), Interim standard for the ion-neutral atomic oxygen collision frequency. *Geophys. Res. Lett.*, **20**, 1543-1546.

Schunk, R. W. and J. C. G. Walker (1973), Theoretical ion densities in the lower ionosphere. *Planet. Space Sci.*, **21**, 1875-1896.

Schunk, R. W. and A. F. Nagy (1978), Electron temperatures in the F region of the ionosphere: Theory and observations. *Rev. Geophys.*, **16**, 355.

Sedgemore, K. J. F., J. W. Wright, P. J. S. Williams, G. O. L. Jones and M. T. Rietveld (1998) Plasma drift estimates from the Dynasonde: comparison with EISCAT measurements. *Ann. Geophysicae* , **16**, 1138-1143.

Shimazaki, T. (1955), *J. Radio Res. Labs., Jpn* **2**, 85.

Sipler, D. P., B. B. Luokkala and M. A. Biondi (1982) Fabry-Perot geterminations of mid-latitude F-region neutral winds and temperatures from 1975 to 1979. *Planet. Space Sci.*, **30**, 1025-1032.

Sipler, D. P., M. E. Hagan, M. E. Zipf and M. A. Biondi (1991), Combined optical and radar wind measurements in the F-region over Millstone Hill. *J. Geophys. Res.*, **96**, 21555-21262.

Solomon, S. C. (1991), Optical Aeronomy. *Rev. Geophys.*, Supplement April 1991, 1089-1109.

Smith, R. W., G. Hernandez, K. Price, G. Fraser, K. C. Clark, W. J. Schulz, S. Smith and M. Clark (1994), The June 1991 thermospheric storm observed in the southern hemisphere. *J. Geophys. Res.*, **99**, 17609-17615.

Smith, R. W. (1998) Vertical winds: a tutorial. *J. Atmos. Solar-Terr. Phys.*, **60**, 1425-1434.

St. Maurice, J. P. and R. W. Schunk (1977), Diffusion and heat flow equations for the mid-latitude topside ionosphere. *Planet. Space Sci.*, **25**, 907-920.

Thayer, J. P., G. Crowley, R. J. Niciejewski, T. L. Killeen, J. Buchau and B. W. Reimisch (1995), Ground-based observations of ion/neutral coupling at Thule and Qanaq, Greenland. *J. Geophys. Res.*, **100**, 12189-12199.

Titheridge, J. E. (1991) Mean meridional winds in the ionosphere at 70^{o}N. *Planet. Space Sci.*, **39**, 657-669.

Titheridge, J. E. (1993) Atmospheric winds calculated from diurnal changes in the mid-latitude ionosphere. *J. Atmos. Terr. Phys.*, **55**, 1637-1659.

Titheridge, J. E. (1995a) The calculation of neutral winds from ionosphere data. *J. Atmos. Terr. Phys.*, **57**, 1015-1036.

Titheridge, J. E. (1995b) Winds in the ionosphere – a review. *J. Atmos. Terr. Phys.*, **57**, 1681-1714.

Thorne, A. P. (1988) *Spectrophysics*. 2nd ed. Chapman and Hall, London.

Vaughan, J. M. (1989) *The Fabry-Perot Interferometer: History, Theory, Practice and Applications*. Adam Hilger series on Optics and Optoelectronics, IOP Publishing Ltd.

Volland, H. (1979). Magnetospheric electric fields and currents and their influence on large-scale thermospheric circulation and composition. *J. Atmos. Terr. Phys.*, **41**, 853-866.

Wayne, R. P. (1985) *Chemistry of atmospheres* , Oxford University Press, Oxford.

Witasse, O., J. Lilensten, C. Lathuillere and B. Pibaret (1998) Meridional thermospheric neutral wind at high latitude over a full solar cycle. *Ann. Geophysicae* , **16**, 1400-1409.

Witasse, O., J. Lilensten, C. Lathuillere and P.-L. Blelly (1999) Modeling the OI 630.0 and 557.7 nm thermospheric dayglow during EISCAT-WINDII coordinated measurements. *J. Geophys. Res.*, **104**, 24639-24655.

Wark, D. Q. and J. M. Stone (1955) Measurement of the Doppler width of the λ 5577 line in the night sky. *Nature*, **175**, 254-255.

Yagi, T. and P. L. Dyson (1985), The response of the mid-latitude thermospheric wind to magnetic activity. *Planet. Space Sci.*, **33**, 461-467.

ACKNOWLEDGEMENTS

First, and foremost, I'd like to thank my family, especially my parents, for their support and love through the years, it is truly appreciated. I have been very lucky to have had a series of supervisors whose patience has only been matched by their good natures and for this I thank Anasuya Aruliah, Alan Aylward and Nigel Meredith and hope it wasn't too trying an experience for them. In an unofficial capacity Ian McWhirter has been as good as any supervisor and I'd like to thank him for his guidance and seemingly inexhaustible patience. I also owe a debt of gratitude to the rest of the team at APL, who have made the last few years at the lab a great working environment, in particular my fellow postgrads Bevis, Matt, Ingo and Yvan.

The nature of the subject of this thesis has meant that I have had an enormous amount of help not just within the lab but also from our collaborators at the remote sites and our colleagues in general. At APL the FPI programme has relied on the dedication and expertise of many people including Jim Percival and Kevin Page. At RAL both Chris Davis and Richard Stamper have made much appreciated contributions. At Kiruna we are indebted to the co-operation of Urban Brandstrom and Ake Steen who have assisted us for many years. I'd like to thank Kent Miller for providing me with a copy of his Meridional Wind Model and for some valuable advice.

I've been lucky enough to make, and keep in contact with, friends in London, Dublin and further afield and I'd like to thank them for their support and putting up with the mood swings. This process has probably gone on too long to actually remember everyone who has helped me and to whom I owe thanks, so if there are any glaring omissions then I apologise.