UNIVERSITY COLLEGE LONDON

Faculty of Mathematics and Physical Sciences

Department of Physics & Astronomy

Forecasting and Modelling the Terrestrial Effects of Space Weather using STEREO and CMAT2

Thesis submitted for the Degree of Doctor of
Philosophy of University College London

by
David Barnes

Supervisors:	Examiners:
Prof. Alan Aylward	Dr. Sarah Matthews
Dr. Chris Scott & Dr. Jackie Davies	Prof. Steve Milan

January 9, 2017

Statement of Authorship

I, David Barnes, confirm that the work presented in this thesis is my own. Where another person's work has been used, this has been acknowledged in the main text of the thesis.

David Barnes

Abstract

The work presented in this thesis focuses on two aspects of space weather. The first of these uses NASA's Solar TErrestrial RElations Observatory (STEREO) to measure solar wind properties. The Heliospheric Imagers (HIs) on-board each of the two STEREO spacecraft are used to infer the density distribution within two Coronal Mass Ejections (CMEs) by tomographic inversion of white light images. The resulting densities from successive images are then used to determine how the CMEs evolve through the heliosphere and to make estimates of their speed and density at Earth, which are compared with measurements from near-Earth spacecraft.

The second area of research involves modelling studies of how the neutral atmosphere responds to energy transfer from the solar wind. Data from the Super Dual Auroral Radar Network (SuperDARN) provide high spatial and temporal resolution measurements of Earth's high latitude electric field in both hemispheres. These data are used to drive the UCL Coupled Middle Atmosphere Thermosphere (CMAT2) General Circulation Model (GCM), which is used to study the effects of electric field variability and hemispheric asymmetry in ionosphere-thermosphere coupling. Results of simulations using this new version of CMAT2 are compared to ground- and space-based observations.

As a means to connect these two areas of research, it was one original aim of this thesis to develop CMAT2 in such a way that STEREO observations could be used to drive its solar wind energy input. Chapter 5 shows that some limited information about geo-effective solar wind properties, such as speed and density, may be determined using HI data. However, detailed information about the interplanetary magnetic field (IMF) is required to accurately simulate energy transfer between the ionosphere and neutral atmosphere. Instead, a new electric field model has been developed, which allows CMAT2 to be driven using the IMF as an input. This model quantifies the electric field variability, as a means to accurately model Joule heating, which is a problem common to many GCMs.

Acknowledgements

I wish to first thank everyone in the APL group at UCL, in addition to the staff at RAL, who have all helped me throughout my PhD. This includes, in particular, my supervisor, Alan, for providing much appreciated guidance since I began at UCL. Anasuya, who has devoted much of her time to helping and encouraging me in Alan's absence. Dr. Chris Scott, who introduced me to the PhD, and who helped me out very early on with my work on STEREO. Dr. Jackie Davies, for taking over Chris' role towards the end of my PhD and for providing encouragement and feedback on this work. Dr. Tim Spain for having the patience to help me with my modelling work when I first began my research. Lastly, my APL colleagues, David and Amy, whom have both been excellent company for the past three years and Rosie for devoting a significant amount of time to proof-reading this thesis.

I would like to also acknowledge those outside UCL and RAL who have assisted with my research in so many ways. Firstly, everyone in the Radio & Space Plasma Physics Group at the University of Leicester, who have provided me with the SuperDARN data and taken the time to help me with any queries relating to it, making this thesis possible. The solar and terrestrial scientists and students, whom I have met at conferences, workshops and meetings, for providing advice and feedback on my work, in addition to great company. Finally, everybody in the astrophysics department who has helped me in any way with my work, from coding issues to thesis writing, I am very grateful.

I must also say thank you to my fellow UCL students, and post-docs, who have made this an incredible department to be a part of, it would not be possible to list all of the fantastic memories I have with you all. Obviously there are too many of you to name personally, but thanks to everyone for making this such a great place to work.

Contents

List of Figures

This page was intentionally left blank

List of Tables

Chapter 1

Overview

Coronal Mass Ejections (CMEs) are large releases of magnetised, energetic particles that originate on the Sun and expand into interplanetary space. Small events occur frequently and, depending on their direction, can reach the Earth within a matter of days. Whilst the Earth's magnetic field generally shields us from the solar wind and CMEs, they are able to transfer energy to the Earth's atmosphere via magnetic reconnection. Solar wind plasma entering the magnetosphere penetrates into the polar regions of the upper atmosphere creating electric fields and producing the aurora. If a very large CME impacts the Earth and possesses the appropriate magnetic orientation it can cause a significant disturbance known as a *geomagnetic storm*.

One such event occurred in September 1859, causing worldwide disruption to telegraph communications and aurorae to be witnessed at latitudes as low as 23°. Some telegraph operators were even able to disconnect power supplies and maintain entire conversations on auroral current alone. Amateur astronomer Richard Carrington, in honour of whom this event is named, had witnessed a large solar flare the previous day. He correctly proposed that these occurrences were connected (Carrington, 1859), giving rise to the field of scientific research, which would become known as *space weather*.

Although the 1859 geomagnetic storm remains the largest such event on record, the actual CME itself was unremarkable. More energetic CMEs have occurred since but caused less severe storms, which is likely to be a result of their magnetic field orientation. The most notable of these events were in August 1972 (Lin and Hudson, 1976) and June 1991

(Kane et al., 1995), where the authors have estimated the respective energies to be 10^{26}J and 10^{27}J. In comparison, Tsurutani et al. (2003) estimate the 1859 CME to have an energy in the region of 10^{25}J. Due to the infrequency of CMEs exceeding this energy it is not possible to statistically infer the probability of such events, however it is clear that the Sun has the potential to cause such a severe storm again. We are able to use less direct means, such as cosmic ray produced isotopes stored in tree rings and ice cores (Steinhilber et al., 2012), to infer the history of solar activity over several thousand years, but this does not reveal detailed information about individual CMEs.

Whilst the Carrington storm was not considered particularly damaging to nineteenth century civilisation, it is easy to imagine how devastating such an event could be today. With modern dependence on electronic technologies and satellite communications, another Carrington storm poses a serious danger. Furthermore, in the coming decades both government space agencies and private spaceflight companies, are planning manned missions to low Earth orbit, the Moon and beyond. CMEs that may pose no risk to humans on Earth present a serious radiation hazard to astronauts, particularly to those who venture beyond the protective magnetic field of the Earth.

It is clear then that, whilst difficult to predict, we can expect a large geomagnetic storm to occur again. The best we can hope for is to find ways to understand their effects and to develop techniques to pre-empt their arrival. The work presented in this thesis explores aspects of both. Chapter 2 presents a summary of the relevant theory and chapter 3 contains descriptions of the relevant instrumentation used, and refered to, throughout this thesis. Next, chapter 4 contains a more in-depth description of the UCL Coupled Middle-Atmosphere Thermosphere (CMAT2) General Circulation Model (GCM). Following this, chapters 5 to 8 present work that has been performed specifically as part of this thesis. Firstly, white-light observations from NASA's Solar TErrestrial RElations Observatory (STEREO) are used to estimate the densities and velocities of Earth-directed CMEs in chapter 5. Secondly, the CMAT2 GCM has been developed to simulate the energy input from the solar wind using data from the Super Dual Auroral Radar Network (SuperDARN) (chapters 6 and 7). Lastly, in chapter 8, the model is developed to be driven using *geo-effective* solar wind properties, such as magnetic field strength and orientation, in order to model the effects of space weather on the upper atmosphere. The final chapter, 9, presents a summary of these findings, and potential future work.

Chapter 2

Theory

2.1 The Terrestrial Atmosphere

2.1.1 Basic Atmospheric Structure

The temperature of the Earth's atmosphere is generally determined by absorption and emission of radiation. Its structure can be described by dividing it into near-spherical shells determined by its temperature gradient (figure 2.1), where different physical processes govern heating and cooling. The boundaries, or *pauses*, between each shell are surfaces of near-constant temperature where the temperature gradient reverses.

The Lower Atmosphere

The lowest layer is the troposphere, where heating is dominated by absorption of IR radiation by water vapour and emission of IR radiation by the Earth's surface. Above the tropopause is the stratosphere where the mean temperature increases from ∼210K at 12km to ∼260K at 45km due to absorption of UV radiation by ozone. In the mesosphere, the mean temperature then drops again towards the mesopause as a result of low absorption and IR emission from CO_2. The mesopause is the coldest region of the atmosphere at ∼180K.

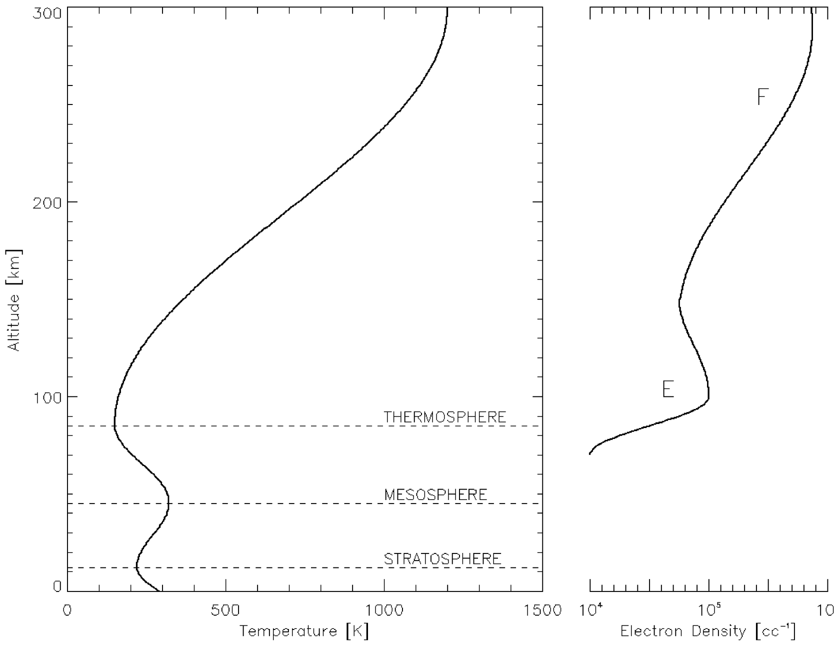


Figure 2.1: The temperature and electron number density profiles of Earth's atmosphere. These represent mean values dependent on many factors including season, time of day and solar activity. The right hand plot indicates the *E* and *F* regions of the ionosphere.

The Thermosphere

Above the mesopause is the thermosphere where absorption of UV and EUV radiation cause the temperature to increase significantly. The temperature here is highly variable due to its strong dependence on solar activity. Above ∼500km is the exosphere, where collisions are so rare that particles follow ballistic trajectories.

The density of the atmosphere decreases sharply as a function of altitude, with more than 90% of the mass lying within the troposphere. The composition of the atmosphere below ∼105km (the homosphere) is about 80% N_2 and 20% O_2 and is subject to turbulent mixing. Above this is the heterosphere where constituents are separated by diffusion into layers determined by their molecular weights. The boundary between these two regions is called the *turbopause*.

The Ionosphere

The strong absorption of UV and EUV in the thermosphere causes ionisation and heating, resulting in a mixture of ions, electrons and neutral molecules, referred to as the ionosphere. The ionosphere is further divided into sub-layers. The F region is a result of atomic oxygen ionisation by EUV radiation and the E region is caused by O_2, N_2+ and NO+ ionisation by soft x-ray and UV radiation. During the day, a further 'D' region is created below the E layer, where NO is ionised by Lyman-alpha radiation. Heating in the ionosphere is also caused by precipitation of solar wind particles in the polar regions and frictional collisions between ions and neutrals.

2.1.2 Vertical Temperature Gradient

Hydrostatic Equilibrium

The density of the Earth's atmosphere decreases with increasing altitude due to the downward gravitational force at a given altitude equalling the upward pressure gradient force. This hydrostatic balance is expressed as

$$\frac{dP}{dz} = -g\rho \tag{2.1}$$

where P is the pressure at altitude z, ρ is mass density and g the acceleration due to gravity. It is a reasonable approximation to assume that g does not vary with altitude, due to the small depth of the Earth's atmosphere when compared to its radius.

The First Law of Thermodynamics

The first law of thermodynamics is the law of the conservation of energy

$$\Delta U = Q - W \tag{2.2}$$

This simply states that the change in energy of a system, ΔU, equals the amount of heat added to the system, Q, minus the work done by it, W. For a compressible fluid this may be written as

$$c_p dT = \frac{1}{\rho} dP \tag{2.3}$$

where T is temperature and c_p the specific heat capacity at constant pressure.

The Adiabatic Lapse Rate

The adiabatic temperature change of a gas results from the heating or cooling it experiences due to compression or expansion, without the transfer of heat with its surroundings. As a parcel of air ascends in the atmosphere the pressure decreases, causing its volume to increase. The work done in driving this expansion comes from a decrease in the temperature of the parcel. Conversely, when a parcel of air is lowered, it experiences an increase in temperature and pressure. The dry adiabatic lapse rate, Γ_D, of the atmosphere is defined as the decrease in temperature with altitude a dry parcel of air will experience when raised or lowered adiabatically. Considering the first law of thermodynamics for an adiabatic process, (2.3) becomes

$$c_p dT = \frac{1}{\rho} dP \tag{2.4}$$

Differentiating with respect to z and substituting (2.1) produces an expression for Γ_D.

$$\Gamma_D = -\frac{dT}{dz} = \frac{g}{c_p} \tag{2.5}$$

where $\Gamma_D \simeq 10^{-3}\mathrm{Km}^{-1}$ in the Earth's atmosphere. Although this would suggest a constant drop-off in temperature with increasing altitude, due to the heating and cooling processes that occur at different layers in the atmosphere the temperature gradient is not perfectly adiabatic.

For a column of air in which the drop in temperature is greater than the adiabatic lapse rate, a parcel of air raised will cool more slowly than its surroundings causing it to rise further. This continuous growth is unstable. The same is true if the parcel of air is lowered adiabatically; it will continue to descend.

Conversely, if the temperature gradient of the column of air is *less* than Γ_D, a parcel of air raised adiabatically will cool faster than its surroundings, causing it to fall again. In the case where the temperature gradient equals the lapse rate, a parcel of air that is displaced vertically will oscillate vertically at the *Brunt-Väisälä* frequency.

2.1.3 Horizontal Winds

Geostrophic Balance

Solar heating in the atmosphere creates a pressure gradient, which will cause a parcel of air to move under its influence. In the rotating reference frame of the Earth, this motion will cause the parcel of air to experience a Coriolis force. These effects will influence the horizontal motion of the parcel of air according to

$$\frac{d\mathbf{V}}{dt} = -\frac{1}{\rho}\nabla p + f\mathbf{V} \times \mathbf{k} \tag{2.6}$$

where $\mathbf{V}$ is the horizontal wind vector, $\mathbf{k}$ the unit vector in the direction of the Earth's spin axis and ∇ the horizontal derivative operator. The Coriolis parameter, f, is a function of the Earth's rotation frequency, Ω, and latitude, θ,

$$f = 2\Omega \sin\theta \tag{2.7}$$

The Coriolis term in (2.6) produces a clockwise force in the northern hemisphere and an anticlockwise force in the southern hemisphere. If this force is equal to the pressure gradient, the system is in *geostrophic balance* resulting in a constant horizontal motion, $\mathbf{V}_g$,

$$\mathbf{V}_g = \frac{1}{f\rho}\mathbf{k} \times \nabla p \tag{2.8}$$

This is a good approximation in the middle atmosphere (the strosphere and mesosphere) where the pressure gradient and Coriolis forces are dominant and is responsible for the circulation of air around regions of high and low pressure. In the mesosphere a global pressure gradient results from the daily variation in solar UV heating, which produces mesospheric jets, as shown in figure 2.2.

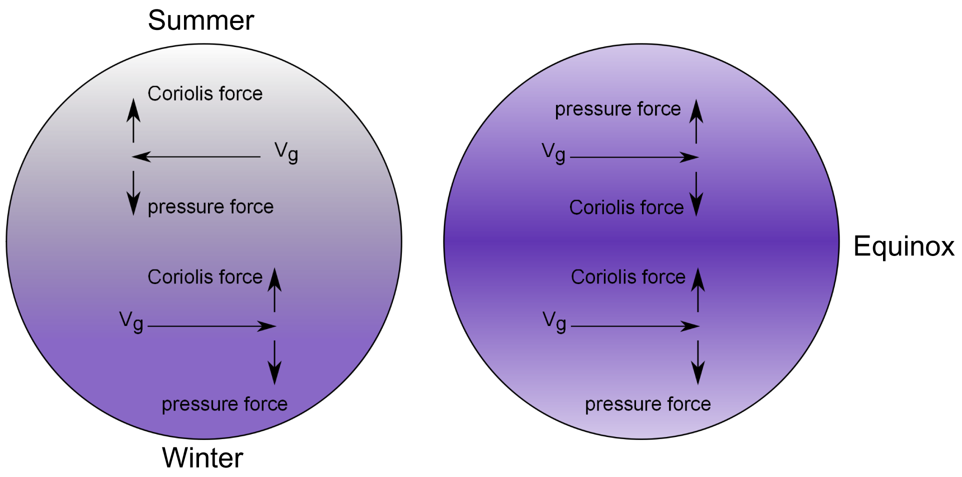


Figure 2.2: Mesospheric zonal wind jets as a function of season.

The Mesopause Anomaly

The temperature structure of the mesopause expected from a perfect geostrophic balance would be hottest in the summer hemisphere and coldest in the winter hemisphere as a result of solar heating. However, observations of the mesopause reveal that the opposite is true, a phenomenon known as the *mesopause anomaly*. Murgatroyd and Singleton (1961) proposed an inter-hemispheric circulation to resolve this problem. An adiabatic upwelling of air in the summer hemisphere results in cooling, whilst a downwelling in the winter hemisphere causes it to be heated. This meridional circulation of air around the mesopause would experience a Coriolis force and so a drag term must be added to (2.6)

$$\frac{d\mathbf{V}}{dt} = -\frac{1}{\rho}\nabla p + f\mathbf{V} \times \mathbf{k} - \mathbf{F} \tag{2.9}$$

For the drag to counteract the Coriolis force, it must act in a westward direction in the winter hemisphere and an eastward direction in the summer hemisphere. Houghton (1978) proposed that drag caused by the dissipation or *breaking* of gravity waves could be responsible for this mechanism. Modelling studies by Lindzen (1981) and Holton (1981), which estimated the momentum dissipation of gravity waves propagating upward from below the mesosphere, were able to demonstrate that they would result in a net acceleration in the required direction.

2.1.4 Gravity Waves

Gravity waves are small scale transverse perturbations in wind, temperature and density, which primarily originate in the troposphere. Sources include the interaction of winds with surface topography, storm convection and wind shear. Their mechanism for dissipating energy can be understood by considering how their amplitude changes as they propagate upward. Conservation of kinetic energy density may be written as

$$E = \frac{1}{2}\rho(z)|\mathbf{V}(z)|^2 = k \tag{2.10}$$

where k is a constant and $|\mathbf{V}|$ is the amplitude of the velocity of the wave. Atmospheric density decreases exponentially with altitude as

$$\rho(z) = \rho_0 e^{-z/H} \tag{2.11}$$

where H is the scale height (section 4.1.1) and ρ_0 is the density at $z = 0$. $|\mathbf{V}(z)|$ may then be expressed as a function of altitude by

$$|\mathbf{V}| = \frac{2k}{\rho_0} e^{z/2H} \tag{2.12}$$

Therefore the amplitude of gravity waves increases exponentially with altitude due to the density drop-off. As the amplitude reaches a critical level, the local temperature gradient produced becomes superadiabatic, causing it to become unstable. The wave begins to break, depositing its energy and momentum into the mean wind flow.

Due to the small-scale size of these waves, relative to the grid sizes used by atmospheric computer models, they cannot be modelled directly. The use of higher resolution models is avoided because instabilities would arise from other sources and the computational costs would be impractical. Rather, the effects of gravity waves are represented by one of several parameterisations.

2.1.5 Solar Tides

Atmospheric tides are periodic oscillations in winds, temperature and pressure that result from either solar heating or the gravitational influence of the Moon[1]. Solar, or *thermal*, tides are produced by radiative heating throughout the atmosphere and dominate over gravitational tides. The heating results from the processes described in section 2.1.1; absorption of IR by water in the troposphere, of UV by O_3 in the stratosphere and of EUV by O in the thermosphere. Due to the conservation of energy, tides originating in the middle and lower atmosphere increase in amplitude as they propagate upward in the same way that gravity waves do. Waves generated below the thermosphere therefore dominate over those generated within it.

Hough Modes

The tides occur with periods that are harmonics of the 24-hour solar day and follow the apparent motion of the Sun. The 24h, 12h and 8h components are termed *diurnal*, *semidiurnal* and *terdiurnal*, respectively. Tidal oscillations are generally found to be dominated by semidiurnal, rather than diurnal tides. This is because the vertically propagating components of the semidiurnal tide are found to be stronger than those of the diurnal tide, causing the latter to be subject to destructive interference (Chapman and Lindzen, 1970). An analytic treatment of the oscillations caused by thermal tides (Lindzen, 1967) uses the tidal equations of Laplace to decompose them into *Hough modes.* These are described using zonal, s, and meridional, n, wavenumbers and are written as (s, n). For the purposes of modelling in CMAT2, only the diurnal ($(1, 1)$) and semidiurnal modes ($(2, 2)$, $(2, 3)$, $(2, 4)$ and $(2, 5)$) are used. Thermal tides have a significant effect on coupling between different levels of the atmosphere and are responsible for the transport of chemical constituents.

2.1.6 Chemistry

Chemical abundances in the atmosphere are governed by photochemical production/loss and transport processes. The combination of these effects is understood by considering the rates at which they each occur.

---

[1] Gravitational tides due to the Sun do also occur, however their magnitude is significantly less than those of the Moon.

Chemical Timescale

The net rate of change of a species due to chemical processes is dependent on all of the possible production and loss mechanisms. For example, for a given species, A, these may be

$$[A] + [B] \rightarrow products(k_1) \tag{2.13}$$

$$[A] + [C] + [M] \rightarrow products(k_2) \tag{2.14}$$

$$[D] + [E] \rightarrow [A] + products(k_3) \tag{2.15}$$

where $[A]$ is the number density of species, A, and k_i the corresponding reaction rate coefficient. The rate of change of species A is given by the sum of these processes

$$\frac{d[A]}{dt} = -k_1[A][B] - k_2[A][C][M] + k_3[D][E] = [A](-L) + G \tag{2.16}$$

where the chemical loss of A, $(-L)$, is dependent on $[A]$, whilst the gain, G, is not. Integrating over time from $t = 0$ gives

$$[A] = \frac{G - (-[A_0]L + G)e^{-Lt}}{L} \tag{2.17}$$

where $[A_0]$ is the value of $[A]$ at $t = 0$. For a process with no chemical production reactions, $[A] = [A_0]/e$ may be substituted into the above equation to define a timescale, τ, over which $[A]$ is reduced by a factor of e

$$\tau = \frac{1}{L} \tag{2.18}$$

Dynamical Timescale

Dynamical production and loss rates can be derived in a similar way, because the concentration of a species falls exponentially with altitude. A *scale height*, H_i, is the vertical distance that results in the decrease of $[A]$ by a factor of e

$$[A] = [A_0]e^{-z/H_i} \tag{2.19}$$

Transport due to vertical wind, w, can then be expressed by differentiating the above equation to give

$$\frac{d[A]}{dt} = -w\frac{d[A]}{dz} = \frac{w}{H_i}[A] \tag{2.20}$$

where $w = -\frac{dz}{dt}$ is the downward wind speed. Therefore, a time constant, H_i/w, is defined over which $[A]$ drops by a factor of e. In a similar way, time constants for horizontal diffusion may be defined by replacing H_i by a scale length.

The dominant process, chemical or dynamic, is then dependent on the relative timescales from the calculations above. If the chemical timescale is much smaller than the dynamical timescale for a given constituent, we may assume *photochemical equilibrium* and the abundance of that constituent can be considered a function of chemical processes alone. Below the turbopause this is the case, where the constituent species are well-mixed. Conversely, dynamic transport dominates in the heterosphere, causing the constituents to separate.

Mean Composition

Figure 2.3 shows the mean composition in the upper atmosphere from the NCAR 1D TIME model and MSIS-E90 (section 3.3.3). The MSIS-E90 model provides neutral temperatures and densities of the Earth's atmosphere, based on averages of ground- and space-based measurements. The major constituents are N_2 and O_2 with approximately constant concentrations of 78% and 21% respectively, due to their long chemical timescales. At higher altitudes, there is a general increase in monatomic constituents, O, N and H due to increased photoionisation and longer chemical timescales.

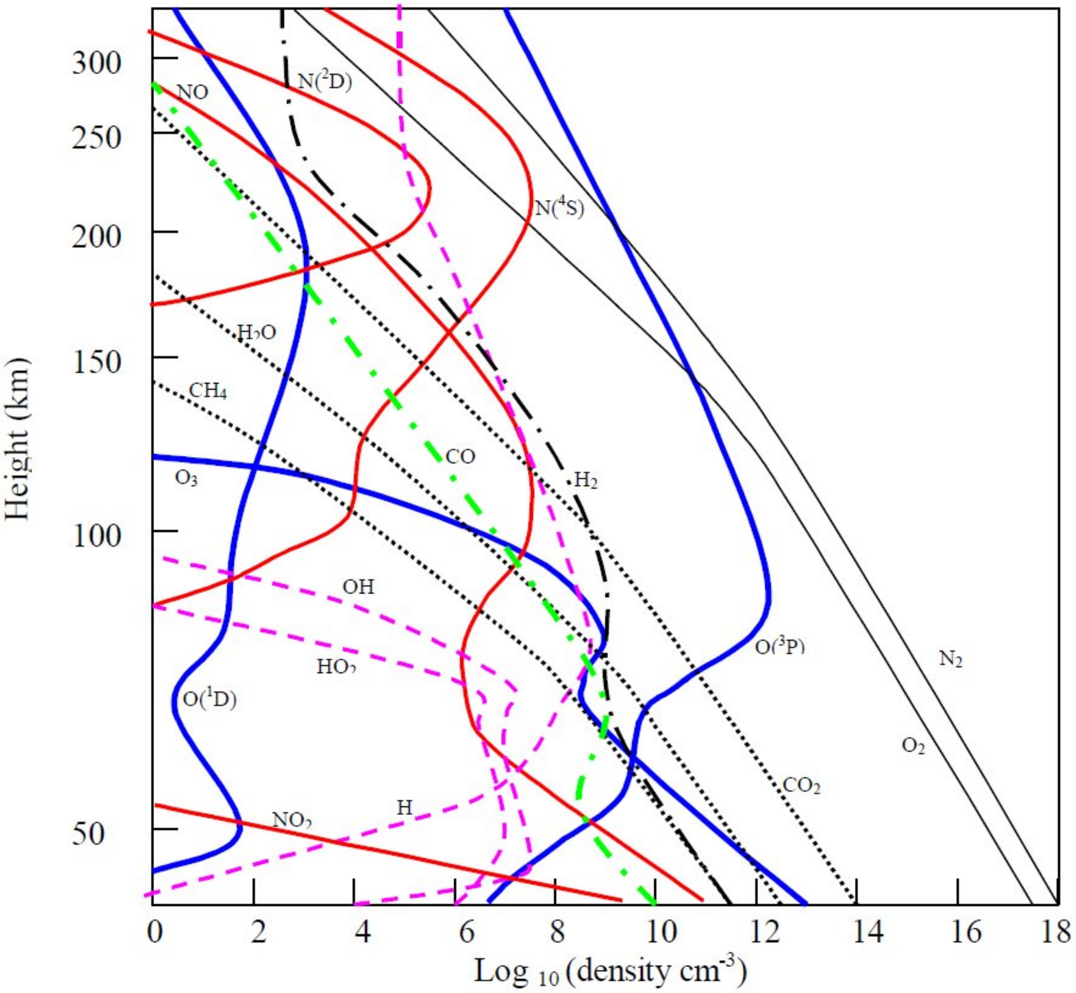


Figure 2.3: Atmospheric composition as a function of altitude based on output from NCAR 1D TIME model (Harris, 2001).

2.2 The Solar Wind

2.2.1 The Sun

The Sun is the dominant source of energy within the Solar System, fusing 6×10^{11}kg of protons into Helium nuclei every second, responsible for an output of 3.8×10^{26}W. It contains more than 99% of the mass of the Solar System and is approximately 75% Hydrogen, 25% Helium and 1% metals. The Sun is a near spherical mass of plasma maintained in hydrostatic equilibrium between its gravity and the outward pressure from fusion in its core. Due to its size, radiation does not reach the surface directly from the core, but instead by a combination of radiative and convective processes that define the Sun's internal structure.

Solar Structure

A small fraction of the energy from fusion will escape as neutrinos, while the majority is carried by photons. Photons radiate energy away from the core by continual absorption and re-emission from ions out to about $0.7R_\odot$, forming the *radiative zone*. Above this level solar plasma cools and becomes opaque to thermal radiation and forms the *convective zone*. Between the radiative and convective zones is a boundary layer approximately $0.03R_\odot$ thick termed the *tachocline* and it is believed that circulating currents within this layer are what produce the Sun's large magnetic field. Heating at the base of the convective zone causes thermal columns to rise to its surface. Above this is the visible surface of the Sun, the *photosphere*, where the convective columns below cause this layer to appear granulated. Observed from Earth, the photosphere has approximately a black-body spectrum, peaking in the visible with a temperature of $\sim$6000K.

The solar atmosphere is made up of four consecutive regions that extend from the surface to beyond the orbit of Pluto, ending in a boundary with the interstellar medium. The composition of the lower layers can be inferred from spectral lines within the Sun's black-body spectrum. The coolest region of the solar atmosphere extends from the surface to about 500km above and is known as the *temperature minimum*. Due to its relatively cool temperature of 4,100K simple molecules are able to form. Above this layer is the 2,000km thick *chromosphere* where the temperature increases with altitude to about 20,000K. At the top of the chromosphere is a thin layer called the *transition region* where the temperature increases rapidly to around 10^6K. The detailed mechanism behind this heating is the subject of continuing research, however at least some is known to result from magnetic reconnection. The uppermost region of the solar atmosphere is the *corona*. This is a continuously expanding stream of plasma that forms the solar wind. The Sun produces a strong magnetic field that is the source of phenomena known collectively as *solar activity*. This field results in differential rotation of plasma at high and low latitudes, with an equatorial rotation rate greater than that at the poles. This causes field lines to become twisted and rise through the photosphere, creating sunspots and the source of CMEs. The heliosphere is the region of magnetised plasma surrounding the Sun, which extends from around $20R_\odot$ to the boundary with interstellar space. Its inner boundary is defined as the surface where the solar wind speed becomes super-Alfvénic. Solar wind plasma carries the Sun's magnetic field outward creating the *Interplanetary Magnetic Field*, or IMF.

Solar Differential Rotation

The differing behaviour in the radiative and convective zones is determined by the relative temperature and pressure gradients they possess, where the former obeys the Schwarzchild criterion (2.21) and the latter does not. This is equivalent to the adiabatic lapse rate (2.5) in the Earth's atmosphere.

$$-\frac{dT}{dz} < \frac{g}{c_p} \tag{2.21}$$

If a fluid element of plasma in the solar interior it will cool and expand. However if the Schwarzchild criterion is obeyed, this will be such that it is less buoyant than its surroundings and it will sink again. Above the tachocline, the negative temperature gradient becomes very large and a fluid element displaced upward will become more buoyant than its surroundings and continue to rise. A further consequence of this is that the solar interior below the tachocline rotates as a rigid body, whilst the convective zone behaves as a fluid. An element of plasma on the tachoclyne near the solar equator therefore has far greater angular momentum than a similar element near the either pole. Conservation of angular momentum is such that, as this element is raised through the convective zone to the solar surface, the Sun rotates at a much greater rate near the equator (25.1 days) than it does at the pole (34.3 days).

Solar Variability

Solar variability refers to changes in frequency and energy of CMEs and other density enhancements, as well as long- and short-term changes in the solar irradiance at Earth. The former is the result of physical changes in the Sun itself, whilst the latter depends also on changes in the Earth's orbital characteristics. Much of this variability manifests itself as periodic changes or *cycles*. The most apparent of these cycles is the quasi-periodic 11 year variation in sunspot number that was discovered by Samuel Heinrich Schwabe during the nineteenth century. This is closely connected to the 22 year Hale cycle where the solar magnetic field reverses polarity every half cycle. The solar magnetic field is known to be responsible for space weather events such as CMEs, and so the frequency of these is closely related to the 11 year cycle. This usually is referred to as *the* solar cycle. The peak in sunspot number during each 11 year cycle, called *solar maximum*, varies between the individual cycles.

Milankovitch cycles refer to the periodic variations in solar irradiance, and its distribution over the Earth, that result from properties of the Earth's orbit. The Earth's axial tilt varies over a 41,000 year interval between about 22.1° and 24.5°, and this axis completes a cycle of precession with respect to the stars in 26,000 years. *Apsidal* precession of the ellipse in which the Earth orbits the Sun takes about 112,000 years. Further hypothetical cycles have been proposed that account for changes between individual sunspot cycles. These mainly involve modulations to the amplitude of the 11 year cycle and typically have periods of hundreds of years or more. These long timescale variations are subject to much speculation and are associated with the debate on climate change. Direct sunspot counts date back only as far as centuries, but indirect evidence from isotopes in ice cores, corresponding to cosmic rays, can give an idea of solar behaviour as far back as several thousand years. The focus of this thesis is on the effects of individual space weather events, rather than long term changes and so variability on timescales longer than the 11 year cycle is not relevant.

2.2.2 The Solar Wind

The solar wind is the extension of the corona into interplanetary space. It is a continuous outflow of plasma driven by the large pressure gradient between the hot base of the corona and the interstellar medium (ISM). The presence of a continuous outflow of plasma from the Sun was first proposed in the 1950s based on observations of comet tails (Biermann, 1951) and the mathematical theory was derived by Parker (1958).

Coronal Expansion

Treating the solar wind as a fully ionised, quasineutral, spherically symmetric plasma and neglecting the effects of the Sun's magnetic field and heat conduction, the equations of continuity, momentum and energy can be expressed as

$$\frac{1}{r^2}\frac{d}{dr}(r^2\rho u) = 0 \tag{2.22}$$

$$\rho v \frac{dv}{dr} + \frac{dp}{dr} + \rho G \frac{M_\odot}{r^2} = 0 \tag{2.23}$$

$$\frac{3}{2} v \frac{dp}{dr} + \frac{5}{2} p \frac{1}{r^2} \frac{d}{dr}(r^2 v) = 0 \tag{2.24}$$

where ρ, p and v are the mass density, pressure and radial velocity, respectively. Parker demonstrated that for a hot corona, of the order 10^6K, the pressure in the outer solar system would exceed that of the gas pressure in the ISM by several orders of magnitude. Parker concluded that it is not possible for the corona to be in hydrostatic equilibrium and therefore he pursued a solution that included non-zero velocity (Parker, 1960). The energy equation gives the following expression for the pressure gradient

$$\frac{dp}{dr} = -\frac{5}{3}\frac{p}{v} - \frac{10}{3}\frac{p}{r} \tag{2.25}$$

This can be substituted in to the momentum equation to give

$$\frac{v^2 - a_s^2}{v}\frac{dv}{dr} = \frac{2a_s^2}{r} - \frac{GM_\odot}{r^2} \tag{2.26}$$

where $a_s = \left(\frac{5p}{3\rho}\right)^{1/2}$ is the speed of sound. If we make the assumption that the corona is isothermal then a_s is constant and the above equation can be integrated to give

$$\frac{1}{2}v^2 - a_s^2 \ln v = 2a_s^2 \ln r + \frac{GM_\odot}{r} + C \tag{2.27}$$

where C is a constant of the integration. There are a number of solutions to this equation, however not all of them are physically meaningful. Based on Biermann's analysis of comet tails and the requirement that $p \rightarrow 0$ towards the ISM at large r, Parker favoured the solution where the solar wind originates at subsonic speeds from the base of the corona and is accelerated as r increases, becoming supersonic at some critical point, r_c. Equation

2.26 appears to present a singularity at the sonic point according to

$$\frac{1}{v}\frac{dv}{dr} = \frac{1}{v^2 - a_s^2}\frac{2a_s^2}{r}\left(1 - \frac{GM_\odot}{2a_s^2 r}\right) \tag{2.28}$$

This problem is avoided by rewriting the above equation in terms of $U(r)$, the gravitational potential of a single proton in the Sun's gravitational field

$$\frac{1}{v}\frac{dv}{dr} = \frac{1}{v^2 - a_s^2}\frac{2a_s^2}{r}\left(1 - \frac{U(r)}{2kT}\right) \tag{2.29}$$

where k is Boltzmann's constant and T is coronal temperature. The singularity is avoided by requiring that, at the critical point, $U(r_c) = 2kT$. This shows that the solar wind becomes supersonic when the thermal kinetic energy of an individual particle equals its gravitational potential energy at the point, r_c,

$$r_c = \frac{3}{5}\frac{m_p g_\odot R_\odot^2}{2kT} \tag{2.30}$$

where m_p is the proton mass and $g_\odot$ is the acceleration due to gravity at the solar surface. Writing (2.27) in terms of the Mach number, $M = v/a_s$, and r_c gives an expression for the velocity in terms of temperature

$$M^2 - \ln M^2 = 1 + 4\ln\left(\frac{r}{r_c}\right) + 4\left(\frac{r_c}{r} - 1\right) \tag{2.31}$$

Although the isothermal assumption is an oversimplification of the true solar wind, it provides a good basis for understanding the physical process behind the expansion of the corona. Figure 2.4 shows the velocity profiles from (2.31) for an isothermal corona at 1 and 2 million Kelvin.

The Interplanetary Magnetic Field

The Maxwell-Faraday law, Ampere's law and Ohm's law can be expressed as

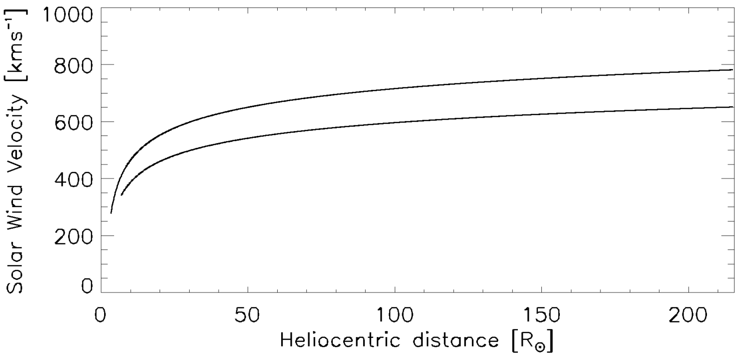


Figure 2.4: Velocity profiles for an isothermal corona, determined from (2.31). The top profile represents 2×10^6K and bottom represents 1×10^6K.

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} \tag{2.32}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{2.33}$$

$$\mathbf{J} = \sigma(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{2.34}$$

where $\mathbf{E}$ is the electric field vector, $\mathbf{B}$ the magnetic field vector, $\mathbf{J}$ the current density and σ electrical conductivity. Substituting (2.32) and (2.33) into (2.34) gives the following

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) - \frac{1}{\mu_0 \sigma} \nabla \times (\nabla \times \mathbf{B}) \tag{2.35}$$

where the relations $\nabla \times (\nabla \times \mathbf{B}) = \nabla(\nabla \cdot \mathbf{B}) - \nabla^2 \mathbf{B}$ and $\nabla \cdot \mathbf{B} = 0$ can be used to derive the induction equation

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) + \frac{1}{\mu_0 \sigma} \nabla^2 \mathbf{B} \tag{2.36}$$

It is reasonable to assume that conductivity in the solar wind is sufficiently high that we can neglect the second right, or advection, term. By Stoke's theorem (2.36) becomes

$$\frac{d}{dt}\int_S \mathbf{B}\cdot d\mathbf{S} = 0 \tag{2.37}$$

where the surface integral can be thought to be over a surface made up of definite fluid elements and the Lagrangian time derivative refers to variation in time whilst following the moving surface elements. This is Alfvén's theorem of *flux freezing* and it implies that the magnetic field moves with the fluid; assuming zero resistivity[2] (infinite conductivity) a fluid element that lies on a magnetic field line will always remain on that field line.

If we assume that the solar wind velocity remains approximately constant above a given heliospheric altitude, R_s, (which is a reasonable assumption based on the conclusions drawn in the previous section) then the position of a stream of particles originating from a given solar longitude, ϕ_0, will obey the equations

$$r - R_s = vt \tag{2.38}$$

$$\phi - \phi_0 = -\Omega_\odot t \tag{2.39}$$

where r and ϕ are heliospheric altitude and longitude respectively, t is time and $\Omega_\odot$ is the Sun's rotational frequency. Combining (2.38) and (2.39) gives

$$\phi(r) = \phi_0 - \frac{\Omega_\odot}{v}(r - R_s) \tag{2.40}$$

which is the equation of an Archimedean spiral. Following from the conclusion that magnetic flux is frozen into the solar wind we can see that the magnetic field lines will also follow this spiral structure, as shown in figure 2.5, commonly known as the *Parker spiral.*

---

[2] This assumption is based the low number density of solar wind particles.

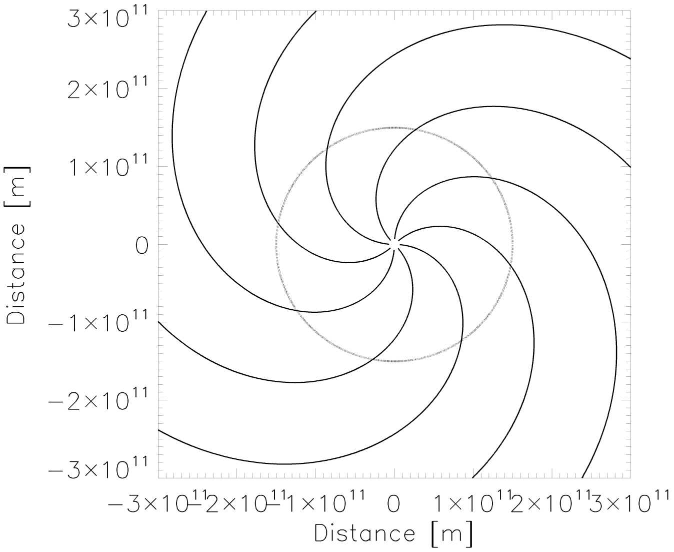


Figure 2.5: Representation of the spiral structure of the IMF produced using (2.40) and assuming a constant velocity equal to $450kms^{-1}$. The dotted circle indicates Earth's orbit.

Heliospheric Current Sheet

The solution derived in section 2.2.2 was based on the approximation that the solar wind is spherically symmetric and does not include the influence of a magnetic field. This is not a representation of its true structure, particularly near the Sun where the majority of particle acceleration is driven by a strong magnetic field. Pneuman and Kopp (1971) derived a solution based on an isotropic corona, but with a steady state, axisymmetric magnetic field and non-uniform $\mathbf{v}$ and ρ. In this case the continuity and momentum equations become

$$\nabla \cdot (\rho \mathbf{v}) = 0 \tag{2.41}$$

$$\rho(\mathbf{v} \cdot \nabla)\mathbf{v} + \nabla p + \rho \frac{GM_\odot}{r^2}\mathbf{e}_r - \mathbf{J} \times \mathbf{B} = 0 \tag{2.42}$$

where $\mathbf{e}_r$ is a unit vector in the radial direction. In the case of a steady state magnetic field, the induction equation and Maxwell-Faraday law become

$$\nabla \times (\mathbf{v} \times \mathbf{B}) = 0 \tag{2.43}$$

$$\mathbf{J} = \frac{1}{\mu_0} \nabla \times \mathbf{B} \tag{2.44}$$

The solution is assumed to be axisymmetric; **B** is a function of radius, r, and spherical polar angle, Θ. Pneuman and Kopp (1971) determined the values of ρ, $\mathbf{v}$, $\mathbf{B}$ and $\mathbf{J}$ via an iterative process by assuming the boundary conditions at the base of the corona. The initial magnetic field is assumed to be dipolar with a surface field strength of 10^{-4}T and an initial velocity, $v_0(\Theta)$, equal to the speed of sound. Starting with the initial dipole field, $\mathbf{J}$ is calculated from (2.44). $\mathbf{B}$ and $\mathbf{J}$ are then used to determine ρ and $\mathbf{v}$ from (2.41) and (2.42). A new magnetic field can then be calculated from (2.43). This process is repeated until the solution converges. Figure 2.6 shows the resulting magnetic field and the initial dipolar field.

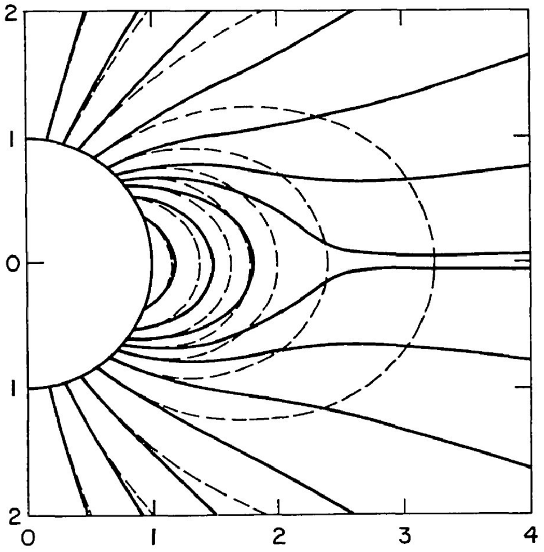


Figure 2.6: Coronal magnetic field line structure, based on a dipole at the base of the corona (Pneuman and Kopp, 1971). The dashed lines represent the initial dipole field used in the derivation.

The resulting structure is composed of closed magnetic field lines at low solar latitudes and open field lines emerging from polar regions. The closed field lines are contained in static equilibrium by the pressure gradient and gravitational forces. The field lines at higher latitudes are stretched outward to large heliocentric distances by the expanding coronal plasma. They do in fact become closed at sufficiently large distances, however these are large enough that they can be considered open. These open field lines are responsible for the solar wind and expand to cover the entire 4π solid angle beyond several $R_{\odot}$.

There are two important consequences of this magnetic field structure that are both apparent from observations of the solar wind and IMF. Firstly, in the regions of open magnetic field lines at high latitudes, the expansion of the solar wind plasma is not restricted by the magnetic field. This results in high speed, low density solar wind. In the equatorial region the solar wind expansion is contained by the magnetic field and so the expansion is slower resulting in higher density plasma. This is supported by coronal images taken during total solar eclipses, where the dense plasma appears as bright regions, or *coronal helmet streamers*, and the sparser regions, or *coronal holes*, appear dark. Solar wind speed measurements from the Ulysses spacecraft, which operated in a heliocentric orbit with a large inclination to the ecliptic, further support this magnetic field structure. They show slow equatorial flows of $\sim 450\text{km}s^{-1}$ and polar flows of $\sim 700\text{km}s^{-1}$.

A second feature resulting from the Pneuman and Kopp (1971) magnetic field model is a region of current in the equatorial plane. The presence of field lines with opposite polarity in this region suggest the existence of a *heliospheric current sheet* circulating around the dipole axis, according to (2.44). Due to the periodic variation in the Sun's magnetic field, the dipole axis can become significantly tilted compared to the rotation axis. In this case the heliospheric current sheet becomes warped, as is seen in figure 2.7, causing it to intersect the Earth's orbit. This crossing can be observed by near-Earth spacecraft, such as ACE and WIND, as a change in magnetic polarity.

When viewed in-situ near the Earth, the solar wind exhibits periodic enhancements in density, velocity and magnetic field as we cross the current sheet (figure 2.8).

This results from the interaction between fast and slow solar wind streams. As the fast solar wind flows catch up with slower flows ahead of them, they are prevented from mixing due to their frozen-in magnetic fields. This causes the fast solar wind streams to become compressed, increasing their density and magnetic field strength (figure 2.9). If these streams persist for a long enough period of time they can pass over the Earth several

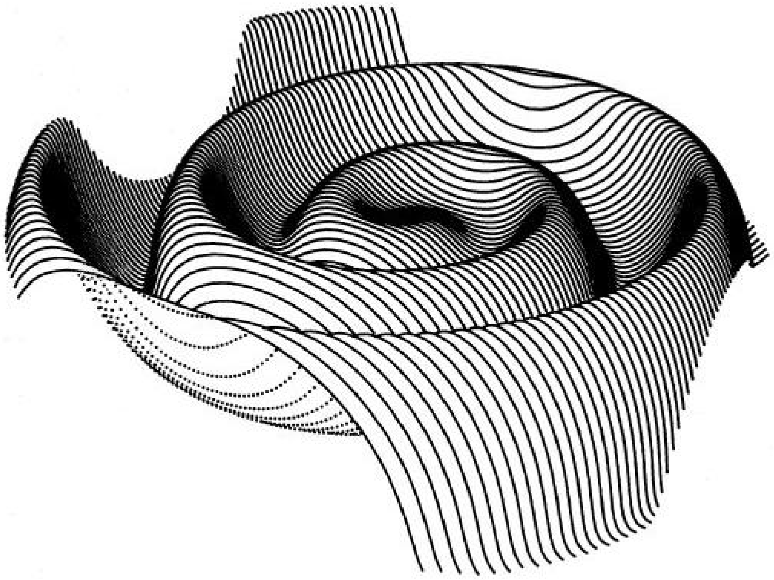


Figure 2.7: Model of the structure of the heliospheric current sheet (Jokipii and Thomas, 1981).

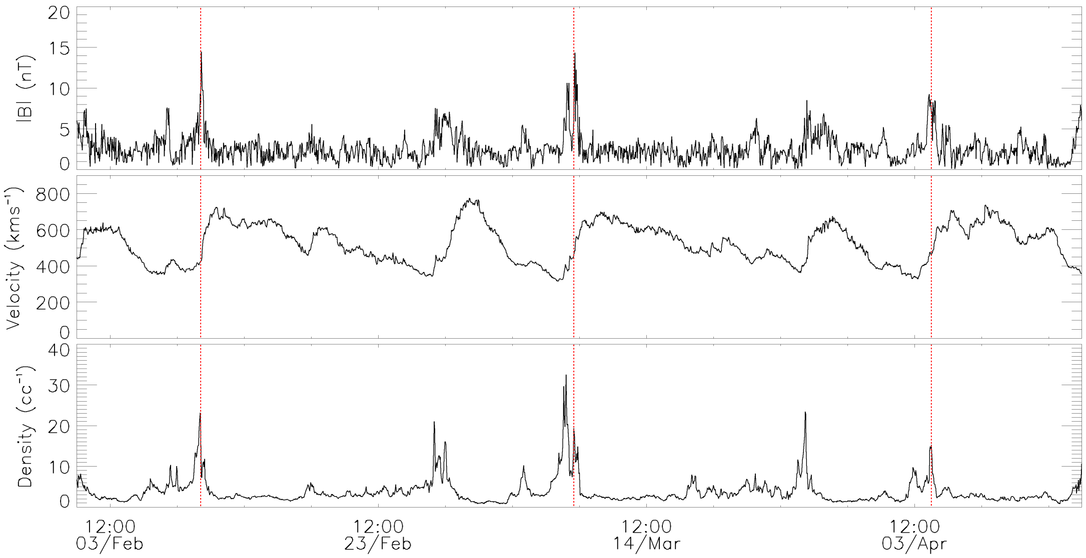


Figure 2.8: IMF magnetic field, velocity and density observed by WIND over several months. The periodic signature of a CIR can be seen every $\sim$ 27 days, indicated by the dashed lines.

times. The Sun rotates with a period of 25.05 sidereal days, resulting in a period relative to the Earth of approximately 26.89 days[3]. A sunward pressure gradient is created between

---

[3] This time difference, Δt, results from the relative sidereal rotation periods of the Sun and Earth, ω_S and ω_E, respectively. The solution is given by $\Delta t = \frac{1}{1-\omega_E/\omega_S}$

the compressed solar wind ahead of the CIR and the less dense region behind it. This acts to slow the CIR slightly, which means the period of a CIR observed at Earth is in fact slightly greater than 26.89 days.

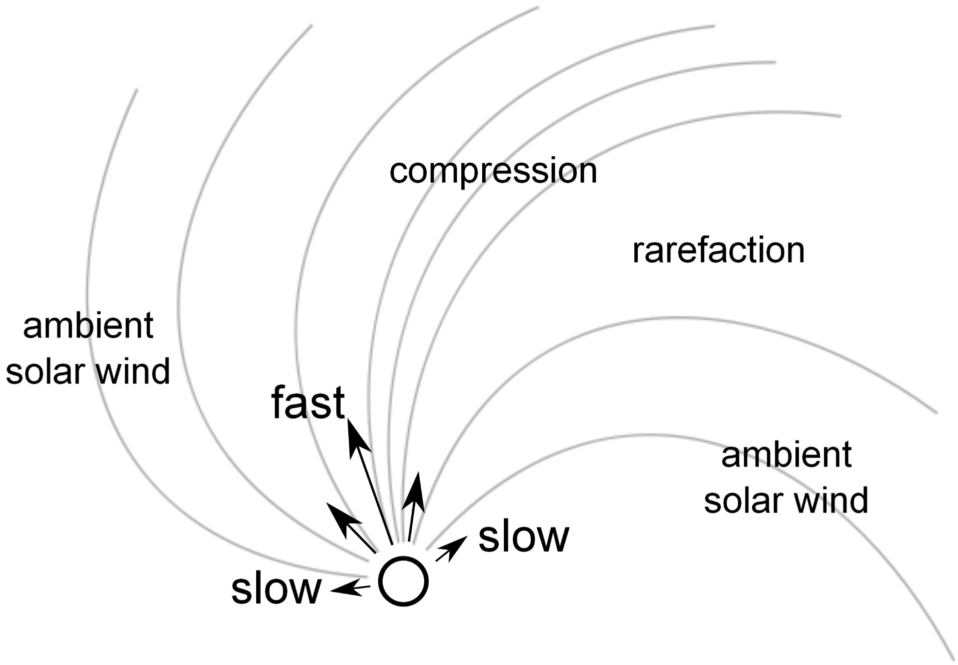


Figure 2.9: Representation of the interaction between fast and slow solar wind streams.

2.3 Coronal Mass Ejections

The most extreme effects of solar activity that are felt on Earth result from the large releases of plasma in CMEs, as discussed in chapter 1. These contain on the order of $\sim 10^{13}$kg of mass and can travel at $\sim 10^{3}\text{km}s^{-1}$. Observations of CMEs date back to the first space-borne white-light coronagraph measurements from Skylab (Gosling et al., 1974) and OSO 7 (Tousey et al., 1974) in the 1970s. The CMEs were seen to be associated with loop-like structures in the EUV corona and, as such, were assumed to be two-dimensional flux tubes (Mouschovias and Poland, 1978) seen in the plane of the sky. Initial attempts at modelling CMEs assumed that they possessed this planar structure, however, evidence quickly emerged that the observed loops were in fact the two dimensional projection of a three-dimensional structure. Howard et al. (1982) studied coronagraph images of an expanding, sun-centred halo. They realised that a CME, resembling an expanding three dimensional bubble, or spherical-shell, would appear in exactly this way, were it viewed head-on.

Understanding of CMEs, with regards to their acceleration mechanisms and structure, has developed significantly from these early, simplistic models. A variety of theoretical models exist to explain CME initiation, however, it is generally accepted that the energy required to power this process comes from the coronal magnetic field. It was originally believed that solar flares were the cause of interplanetary disturbances such as geomegnetic storms and, when CMEs were discovered in the 1970s, it was thought that they were flare-driven. Gosling (1993) dispelled this belief and showed that CMEs are separate phenomena and that it is they that are responsible for geomagnetic disturbances. It is now established that whilst CMEs and flares are often associated, there is not a one-to-one correspondence and that each may be observed in the absence of the other. Instead, flares and CMEs are two parts of a single magnetically-driven process and therefore CME initiation models must be able to account for both phenomena. The most prominent of these, proposed by van Ballegooijen and Martens (1989) and simulated numerically by Amari et al. (2003), is known as the Flux Cancellation model. In this model a magnetic arcade experiences shearing due to external pressure. This causes the convergence of oppositely directed magnetic field lines, which experience reconnection. The resulting magnetic field lines possess a helical structure, which forms a magnetic flux rope containing cool, dense plasma at its core. This is observed as a prominence or filament, depending on its position on the solar disk. Further reconnection can cause this flux rope to become liberated from the coronal magnetic field and it is ejected due to the conversion of magnetic energy to kinetic enrgy. Energy freed by the reconnection is also transported downward, which may responsible for the flare.

CMEs typically possess a three-part structure that is characterised by a bright frontal loop, a dark cavity and a bright core (Webb and Hundhausen, 1987). The frontal loop contains dense plasma swept up by the CME as it propagates through the ambient solar wind, whilst the the cavity appears darker due to its low density but contains the large magnetic field. At the centre of the cavity, the bright core contains dense plasma associated with the erupting filament (House et al., 1981). An individual CME may not exhibit all three of these features, or may exhibit a more complex structure. In fact less than 30% of the SMM and Skylab CMEs studied by Webb and Hundhausen (1987) show all three of these features. CMEs are generally observed to expand in a self-similar manner and, as a result, their observed size is approximately proportional to the distance at which they are observed from the Sun (Low, 2001).

The Large Angle and Spectrometric Coronagraph (LASCO, Brueckner et al. (1995)) has provided coverage of CMEs since it began operation in 1995. Prior to this, the Solwind (1979-89, Sheeley et al. (1980)) and Solar Maximum Mission (SMM, 1980-89, MacQueen et al. (1980)) provided space-based CME observations and, during the period in between (1989-95), ground-based observations, such as the Mauna Loa MK3 Coronameter (Fisher et al., 1981), account for continuous coverage over four solar cycles. Many of the general CME properties were established by Howard et al. (1985) using Solwind observations during the maximum of cycle 21. CME mean speeds were determined to be $450\mathrm{kms}^{-1}$, their angular extend was found to be 45°, centered broadly on the equator, and the total CME rate was found to be $1.8\mathrm{day}^{-1}$, on average. Solar cycle trends were studied by Burkepile and St. Cyr (1993) using the SMM coronagraph, who found that the broad latitudinal distribution disappeared toward solar minimum in 1986, where CMEs were confined to equatorial latitudes. Towards solar maximum of cycle 22 (1989), the broad distribution returned. The launch of SOHO has meant that coverage of CMEs has been provided by the LASCO instruments encompassing solar cycles 23 and 24. These instruments have a combined range from 1.1 to $32\mathrm{R}_\odot$. Yashiro et al. (2004) provide a statistical analysis of nearly 7000 LASCO CMEs during 1996-02, which show an increase in width from 47° (solar minimum, 1996) to 61° (solar maximum, 2002). CME speeds exhibit a similar trend, increasing from approximately $300\mathrm{kms}^{-1}$ at solar minimum to $500\mathrm{kms}^{-1}$ at solar maximum. CMEs are found to approach a constant speed as they reach larger solar distances, however, the slower events exhibit an acceleration and the faster ones a deceleration (Yashiro et al. (2004); Gopalswamy et al. (2009)). This is believed to result from drag forces acting on CMEs, which cause them to converge toward the ambient solar wind speed. Yashiro et al. (2004) show that, by the edge of the LASCO FOV (approximately 4° into the HI-1 FOV), 87% of CMEs have reached a constant speed.

2.3.1 Coronal Tomography

Knowledge of the state of the Earth-directed solar wind in the inner heliosphere is essential to forecasting space weather. Whilst solar wind monitors, such as ACE and WIND, at the stable L1 point provide us with accurate in situ measurements of geo-effective solar wind parameters, their location means that they are only able to give about one hour's advance warning. Spacecraft positioned closer to the Sun than L1 by making use of an electronic solar sail have been proposed (Janhunen, 2010), however they would only improve forecast times by about two-fold. An alternative, then, is to attempt to infer solar wind parameters from afar using photospheric light that has been scattered by electrons in the solar wind, potentially providing several days advance warning of Earth-directed CMEs.

Excluding the background star field, dust and bodies in the solar system, the white light observed by the STEREO heliospheric imagers comes from photospheric radiation that has been Thomson scattered by coronal electrons. The light collected by each CCD pixel in HI is then the sum of all the radiation scattered within the line-of-sight volume. The white light observations made by HI are therefore a function of the electron density distribution along each pixel's line-of-sight. Ambiguities about the true depth of density enhancements arise because the CCD array gives a two-dimensional projection of the optically thin heliosphere. The different vantage points provided by STEREO allow these projections to be combined in order to infer the three-dimensional distribution of electrons, via tomography.

Tomography is the reconstruction of three-dimensional structure from a set of two-dimensional projections. The theory of tomography originated in the medical profession in the early 20^{th} century, motivated by the need for 3D imaging and made possible by the recent discovery of x-rays. The basic principles are equally applicable to the heliosphere, although the technique is very different in practice. Despite the single viewpoint available from near-Earth spacecraft, there were early attempts at tomographic reconstruction of CMEs. These are discussed briefly in order to provide a context for STEREO observations.

Solar Rotational Tomography (SRT)

As the solar surface completes one rotation relative to the Earth every $\sim$26 days, any coronal structures that do not vary significantly on a timescale of less than a few days may be viewed from a range of different angles by an observer on (or near) Earth. This provides the multiple viewpoints necessary for tomography. This is the principle of SRT and was first applied to white-light coronagraph images from Skylab (Altschuler, 1979). Further attempts were made by Zidowitz et al. (1996); Zidowitz (1997, 1999) using both the Mauna Loa MK3 Coronameter and the inner LASCO coronagraph, C1, onboard SOHO. (Frazin, 2000; Frazin and Janzen, 2002) applied the technique to LASCO C2, which extends further out to $6R_\odot$. Jackson et al. (2003) combined ground-based interplanetary scintillation (IPS) observations with data from the Solar Mass Ejection Imager onboard Coriolis to apply this technique to corotating structures much further out in the heliosphere.

Polarimetric Reconstruction

Another single viewpoint approach makes use of polarised brightness observations. Due to the nature of the Thomson scattering process (section 2.3.3), the light emitted by an electron is polarised in the plane perpendicular to the direction of the Sun. The ratio of polarised to unpolarised light, or the *degree* of polarisation, can then be used to infer the approximate depth at which any density enhancements are present. This approach was applied by Moran and Davila (2004) to two CMEs, again using LASCO, although line-of-sight effects were still present. In practice this method is often combined with SRT.

Tomography using the Helios Spacecraft

Jackson and Froehling (1995) reconstructed the 3D density of a CME using the zodiacal light photometers on board the two Helios spacecraft, combined with the (near-Earth) Solwind coronagraph. Whilst the HELIOS photometers were designed for studying dust, and therefore not ideal for CME observations, this was the approach to tomography in the corona using two widely spaced viewpoints.

2.3.2 Tomography in the STEREO Era

The means to perform solar tomography have been greatly improved by the launch of STEREO due to the unique viewpoints provided by it. The first attempt to reconstruct cornal electron densities with STEREO instrumentation was performed by Kramar et al. (2009) and used COR1 data. The study uses polarimetric reconstruction and SRT over two 14-day periods, including Carrington Rotation 2066 (CR2066), whereby a static coronal density between $1.5R_\odot$ and $4R_\odot$ is reconstructed using appoximately 1-3 images each day, which are deemed to be “stable”. Butala et al. (2010) apply a dynamic method to two 14-day periods, again including CR2066. This method allows for the expansion of the corona over time, the solution of which is constrained by physical processes. de Patoul et al. (2013) combine SOHO/EIT and both STEREO/EUVI to reconstruct static coronal densities in the lower region between $1.01R_\odot$ and $1.39R_\odot$. Applying these SRT methods to CMEs becomes very difficult because they move through images on a timescale far smaller than a Carrington rotation. Frazin et al. (2009) combine the two COR1 viewpoints with LASCO and apply prior assumptions about the position of the CME surface. The densities are then solved via an iterative process where the position of the CME surface is allowed to evolve such that it better conforms to the data. Frazin (2012) uses the same observations but employs magnetohydrodynamic (MHD) simulations to also inform the reconstruction algorithm. More recently, studies by Kramar et al. (2014) use a combination of COR1 and EUVI observations to retreive electron densities from CR2066 and associate them with open/closed coronal magnetic fields using a 3D MHD (Miki et al., 1999) model of the corona. Kramar et al. (2016) use a vector tomography technique to determine 3D coronal magnetic fields by means of Hanle effect polarisation measurements by the Coronal Multichannel Polarimeter (CoMP) and 3D electron densities from EUVI tomography. These are found to be consistent with the same MHD model used in the previous study, however there exists no means of independent verification.

As is apparent from the majority of work detailed above, these techniques are predominantly applied to coronagraph observations, rather than heliospheric imager data. The typical method for making predictions of the solar wind, or CME impacts, at Earth is to use coronagraph or other near-Sun observations and propagate solar wind plasma by means of an MHD model, most commonly WSA-ENLIL (section 3.1.5). Wide-angle imagers such as HI present an advantage over coronagraphs because solar wind and CMEs may be observed much further out into the heliosphere. However, the ratio of observed brightness from electron- and dust-scatterd light (the K- and F-coronae) drops rapidly with distance from the Sun. Since the K-corona is the objective of solar tomography, the technique is more easily applied to coronagraph images, where it is dominant. Towards the outer edge of the HI-1 fields of view the ratio of observed K-coronal brightness to that of the F-corona is approximately 0.01. The technique becomes more challenging using HI but, combined with appropriate background subtraction, may still be applied. These techniques and the resulting density estimates are the subject of chapter 5.

The unique viewpoints provided by the STEREO mission allow remote sensing estimates of density structures in the solar wind between the Sun and the Earth. By combining white light observations from the Heliospheric Imagers on board each of the STEREO spacecraft it is possible to estimate the three-dimensional electron density distribution via tomographic reconstruction. Two viewpoints are insufficient to infer detailed CME morphology using tomography, however, when used to estimate density distributions it is possible to arrive at a physically consistent result. CMEs are identified using the heliospheric imagers and densities are then estimated by fitting integrated line-of-sight intensities, measured by each spacecraft, to a 3D heliocentric grid based on Thomson scattering of photospheric light by solar wind electrons.

2.3.3 Remote Sensing CMEs using White-Light

Scattering of photospheric light in the heliosphere is via Thomson scattering by free electrons. Minnaert (1930) was the first to derive a quantitative expression for the light scattered within a line-of-sight volume and to include the effects of solar limb darkening. This work was later developed by van de Hulst (1950) and Billings (1967), and a review of these papers has been presented more recently by Tappin and Howard (2009a), where the notation has been modified for the purpose of clarity. This section follows the derivation of Tappin and Howard (2009a), which describes the nature of the Thomson scattering process, how it is applied to the heliosphere and how this relates to STEREO observations.

Thomson Scattering Theory

Thomson scattering is the elastic scattering of electromagnetic radiation by a charged particle. The theory may be applied provided two conditions are met; the wavelength of the light is much less than the particle separation and the photon energy is much less than the rest mass energy of the scattering particle. Both are true of white light scattering in the heliosphere. Tappin and Howard (2009a) present a pictorial summary of the Thomson scattering process from Jackson (1975), with emphasis on how the observed intensity is affected by the angle of scattering, χ (figure 2.10).

The electric field of a single photon is perpendicular to its direction of propagation. When the photon is absorbed by an electron, the electron will be accelerated in the direction of this electric field. The electron will then re-emit a photon in a direction perpendicular to its displacement. The scattered intensity of photospheric light as a function of observing angle χ can be understood by extending this concept to an electromagnetic wave incident on an electron. An observer at $\chi = 0°$ or $180°$ will see the electron displaced equally in all directions in the plane of the sky and so scattered light will be unpolarised. An observer at $\chi = 90°$ will only see linear displacement of the electron, transverse to the direction of the incident radiation. The scattered light seen will therefore be entirely polarised in the transverse direction. For an observer at any other angle, the scattered light will appear

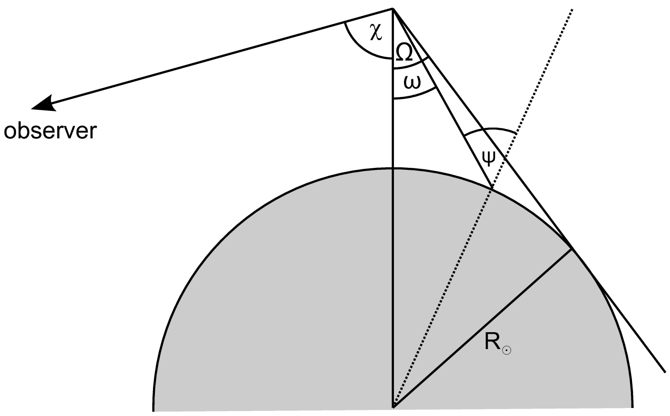


Figure 2.10: The angles relevant to Thomson scattering from a point near the Sun. Modified from Fig. 3 in Tappin and Howard (2009a).

partially polarised. The transverse component of the electric field is independent of χ, whilst the component parallel to the *projected* beam direction will vary with $\cos\chi$. The intensity of the parallel component is then dependent on $\cos^2\chi$. The differential cross-section according to Jackson (1975) is

$$\frac{d\sigma}{d\omega} = \left(\frac{e^2}{4\pi\epsilon_0 m_e c^2}\right)^2 \left(\frac{1+\cos^2\chi}{2}\right) \tag{2.45}$$

where $d\omega$ is the element of solid angle covered by the Sun, at the scattering location. The term in the right hand brackets of (2.45) accounts for the polarisation effects, whilst the term in the left hand brackets is the differential cross section for perpendicular scattering, σ_e,

$$\sigma_e = \frac{e^4}{(4\pi\epsilon_0)^2 m_e^2 c^4} = r_e^2 = 7.95 \times 10^{-30} m^2 sr^{-1} \tag{2.46}$$

Application to the Heliosphere

When applying Thomson scattering theory to white light observations, the theoretical considerations discussed thus far must be extended to account for the finite angular size of the photosphere and for the fact that observed intensities are the sum of contributions from *all* electrons along a given line-of-sight. The first condition is satisfied by integrating each component of the scattered light over the disk of the Sun and the second is met by integrating over the volume contained by the line-of-sight.

Integration over the solar disk, accounting for each of the polarised components, was first performed explicitly by Billings (1967). A review of this is provided by Tappin and Howard (2009a), in which the authors elaborate upon the more difficult concepts in order to avoid any confusion they feel resulted from the original derivation. The lines-of-sight associated with HI observations pass sufficiently far from the Sun that the integration over the solar disk will not be covered in detail here, rather a summary of the important results from Tappin and Howard (2009a) is presented to provide a context for observations away from the Sun.

Integration over the Solar Disk

The tangential and parallel components of the intensity seen by an observer at a distance z from the scattering point are expressed as

$$I_T = \frac{\pi \sigma_e}{2z^2} \int_{\cos\Omega}^{1} I(1 + \cos^2\omega) d(\cos\omega) \tag{2.47}$$

$$I_P = -\frac{\pi \sigma_e}{2z^2} \int_{\cos\Omega}^{1} I \sin^2\chi (1 - 3\cos^2\omega) d(\cos\omega) \tag{2.48}$$

where the intensity is integrated over the angular extent of the photosphere, Ω, (figure 2.10). Solar limb darkening is then included using the expression

$$I = I_0(1 - u + u\cos\psi) \tag{2.49}$$

where u is the wavelength-dependent limb darkening coefficient. $\cos\psi$ may be expressed in terms of the angles ω and Ω, according to the identity

$$\sin\psi = \frac{\sin\omega}{\sin\Omega} \tag{2.50}$$

Substituting (2.49) in to (2.47) and (2.48) yields

$$I_T = I_0 \frac{\pi \sigma_e}{2z^2}[(1-u)C + uD] \tag{2.51}$$

$$I_P = I_0 \frac{\pi \sigma_e}{2z^2} \sin^2 \chi[(1-u)A + uB] \tag{2.52}$$

where A, B, C and D are integrals, of which the van de Hulst coefficients are the solutions

$$A = \cos\Omega \sin^2\Omega \tag{2.53}$$

$$B = -\frac{1}{8}\left[1 - 3\sin^2\Omega - \frac{\cos^2\Omega}{\sin\Omega}(1 + 3\sin^2\Omega)\ln\left(\frac{1+\sin\Omega}{\cos\Omega}\right)\right] \tag{2.54}$$

$$C = \frac{4}{3} - \cos\Omega - \frac{\cos^3\Omega}{3} \tag{2.55}$$

$$D = \frac{1}{8}\left[5 + \sin^2\Omega - \frac{\cos^2\Omega}{\sin\Omega}(5 - \sin^2\Omega)\ln\left(\frac{1+\sin\Omega}{\cos\Omega}\right)\right] \tag{2.56}$$

Approximation away from the Sun

Moving away from the Sun, the integration over the photosphere becomes less significant and the van de Hulst coefficients tend towards the values

$$A = \left(\frac{R_\odot}{r}\right)^2, B = \frac{2}{3}\left(\frac{R_\odot}{r}\right)^2, C = \left(\frac{R_\odot}{r}\right)^2, D = \frac{2}{3}\left(\frac{R_\odot}{r}\right)^2 \tag{2.57}$$

where r is the distance from the Sun to the scattering point (figure 2.11). The total scattered intensity is expressed in terms of the tangential and parallel components by

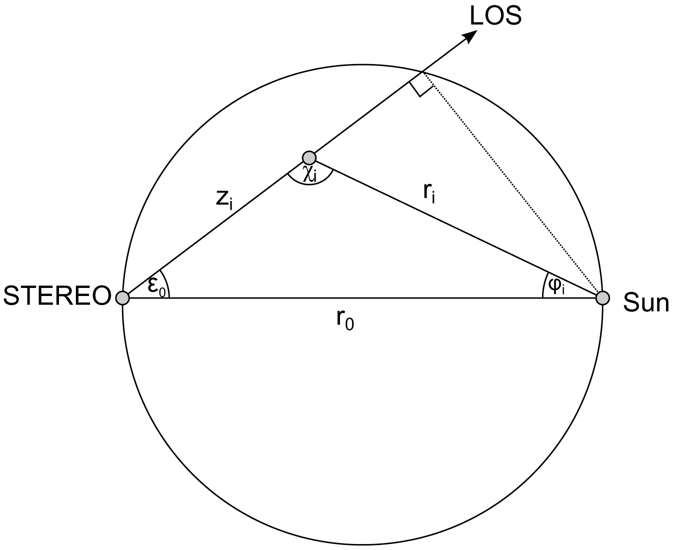


Figure 2.11: The geometry associated with Thomson scattering from a point, i, along a given line-of-sight. The circle represents the *Thomson sphere*, where the differential cross section for scattering (2.45) is maximised.

$$I_{tot} = 2I_T - I_P \tag{2.58}$$

Substituting the approximations for A, B, C and D from (2.57) into (2.51) and (2.52), and the result into (2.58), yields an expression for the total scattered intensity observed from a single electron at a distance z

$$I_{tot} = I_0 \frac{\pi\sigma_e}{2z^2}\left(\frac{R_\odot}{r}\right)^2\left(1-\frac{u}{3}\right)(2-\sin^2\chi) \tag{2.59}$$

Line-of-Sight Integration

The final step in quantifying the observed intensities is to integrate (2.59) over all electrons encompassed by the line-of-sight of the detector. Tappin and Howard (2009a) point out that this is in fact an integration over the point spread function of the observing instrument, rather than a line-of-sight integral. They define a coordinate system where the z axis is along the line-of-sight and the x and y directions mutually orthogonal to z. A volume element of plasma within the detectors field of view will contribute a power of

$$dP = I_{tot}(z)N_e(x,y,z)dxdydz\partial A \tag{2.60}$$

where N_e is the electron number density and ∂A is the surface area of the detector. The power contributed to the detector between a distance z and $z+dz$ is found by integrating (2.60) over the angular extent of the field of view, $\partial\omega$,

$$P_{rec}(z)dz = \int\int_{\partial\omega} I_{tot}N_e dxdydz\partial A \tag{2.61}$$

The following identity may be used to express $d\omega$ at a distance z

$$z^2\partial\omega = dxdy \tag{2.62}$$

This allows (2.61) to be expressed more simply by

$$P_{rec}(z)dz = I_{tot}N_e z^2 dz\partial\omega\partial A \tag{2.63}$$

By setting $I_0\partial A\partial\omega$ to 1, the intensity received by the observer from the element between z and $z+dz$ can be expressed in solar brightness units, $B_\odot$,

$$I_{rec}(z)dz = \frac{I_{tot}}{I_0}N_e z^2 dz \tag{2.64}$$

The total intensity received is then simply found by integrating (2.64) along the line-of-sight

$$I_{rec} = \int_0^\infty \frac{I_{tot}}{I_0}N_e z^2 dz \tag{2.65}$$

It should be noted here that two simplifying assumptions have been used in this derivation. Firstly, electron density, N_e, is sufficiently low that multiple scattering may be ignored and, secondly, the angular coverage of each pixel is sufficiently small that N_e is considered to be constant across the x and y directions.

In order to formulate the observations in terms of an equation that can be solved for densities a grid must be defined over which the densities are to be evaluated. Due to the dependence of (2.65) on r and χ, a spherical, heliocentric grid is chosen (see Appenix A). This means expressing the integral in terms of a new coordinate system by employing the following identities

$$\sin^2\chi \equiv \sin^2(\phi + \epsilon) \tag{2.66}$$

$$r = r_0 \frac{\sin\epsilon}{\sin(\phi + \epsilon)} \tag{2.67}$$

$$dz = r_0 \frac{\sin\epsilon}{\sin^2(\phi + \epsilon)} d\phi \tag{2.68}$$

These are derived from figure 2.11 using basic trigonometric identities. (2.65) then becomes

$$I_{rec} = \frac{\pi\sigma_e R_\odot^2}{2r_0 \sin\epsilon} \left(1 - \frac{u}{3}\right) \int_0^{\pi-\epsilon} (2 - \sin^2(\phi + \epsilon)) N_e d\phi \tag{2.69}$$

where the upper limit, $\pi - \epsilon$, is the limit of ϕ as $z \to \infty$.

2.4 Solar-Terrestrial Energy Transfer

2.4.1 The Magnetosphere

The magnetosphere is the region of space surrounding the Earth, where the planet's own magnetic field strength dominates over the IMF, creating a cavity bounded by the *magnetopause*. Due to the frozen-in magnetic field, supported by the fast moving solar wind, the magnetosphere takes on a non-dipolar shape that is elongated in the anti-sunward direction. Closed magnetic field lines in the dayside[4] magnetosphere are compressed by solar wind ram pressure, forming a bow shock. Those on the nightside become elongated due to the high velocity of the solar wind (figure 2.12). The position of the magnetopause may vary significantly as a result of solar wind variability, however its typical distance on the dayside is approximately 9 Earth radii.

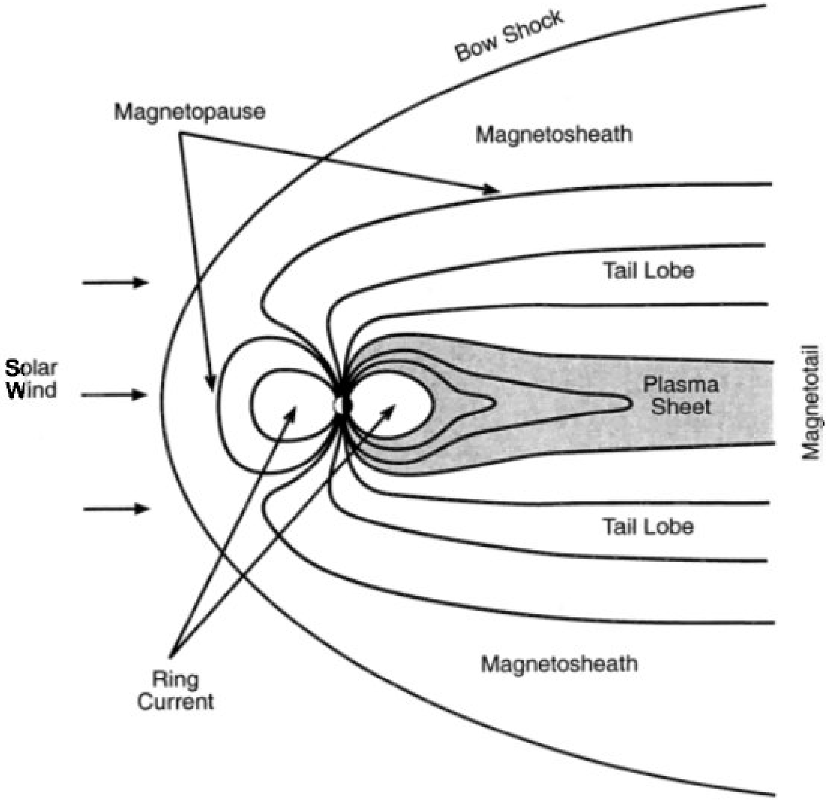


Figure 2.12: Illustration of the main features of the magnetosphere, from Lui (2001).

The elongated nightside of the magnetosphere, referred to as the *magnetotail*, contains north and south *lobes* containing oppositely directed magnetic flux. These lobes have low particle densities of, typically, $10^{-2}\mathrm{cc}^{-1}$, however they are separated by a much denser region called the *plasma sheet*. This is caused by particles in the magnetotail experiencing an $\mathbf{E}\times\mathbf{B}$ drift, which is directed toward the equator. The opposing magnetic field polarities above and below the plasma sheet result in a dawn to dusk current flowing across it, which is closed on the boundary of the magnetotail. Nearer the Earth, the geomagnetic field is approximately dipolar in shape. Because the magnetic field becomes weaker away from

---

[4] Dayside and nightside are terms used to refer to the respective hemispheres of Earth facing towards and away from the Sun. Likewise, *dawn* and *dusk* are used to refer to the respective hemispheres that are emerging from and entering the nightside.

the Earth, charge-dependent $\nabla\mathbf{B}$ and curvature drifts cause a *ring current* to flow in an east to west direction throughout this region of plasma. As solar wind plasma enters the Earth's magnetosphere, crossing the magnetopause, protons and electrons experience opposing forces under the influence of the magnetic field

$$\mathbf{F} = q\mathbf{v} \times \mathbf{B} \tag{2.70}$$

This Lorentz force causes charged particles to gyrate around magnetic field lines, where the *gyro radius* and *gyro frequency* are given by

$$r_i = \frac{m_i|\mathbf{v}_\perp|}{q_i|\mathbf{B}|} \tag{2.71}$$

$$\omega_i = \frac{q_i|\mathbf{B}|}{m_i} \tag{2.72}$$

where q is the charge of the particle and m its mass. As the solar wind particles initially enter the geomagnetic field they complete approximately one half orbit before re-entering the solar wind stream. This causes protons and electrons to drift in opposite directions, causing a current to flow in an eastward direction around the magnetopause. This is termed the *Chapman-Ferraro* current.

The Dungey Cycle

Due to the dependence of the Sun-Earth interaction on the direction of the IMF and solar wind speed and density, perturbations in these values, such as CMEs, can result in large amounts of energy being transferred in a short period. These are divided into two classes; *substorms*, which affect high latitudes and last for a period of hours, and *storms*, which are a global phenomenon and can typically last much longer.

The complex mechanisms that cause substorms are an ongoing area of research and are not fully understood, however there is a general consensus on the energy transfer processes that occur. Akasofu (2004) provides a detailed overview of the findings made in substorm research and how they can be integrated to build up a picture of how substorms occur.

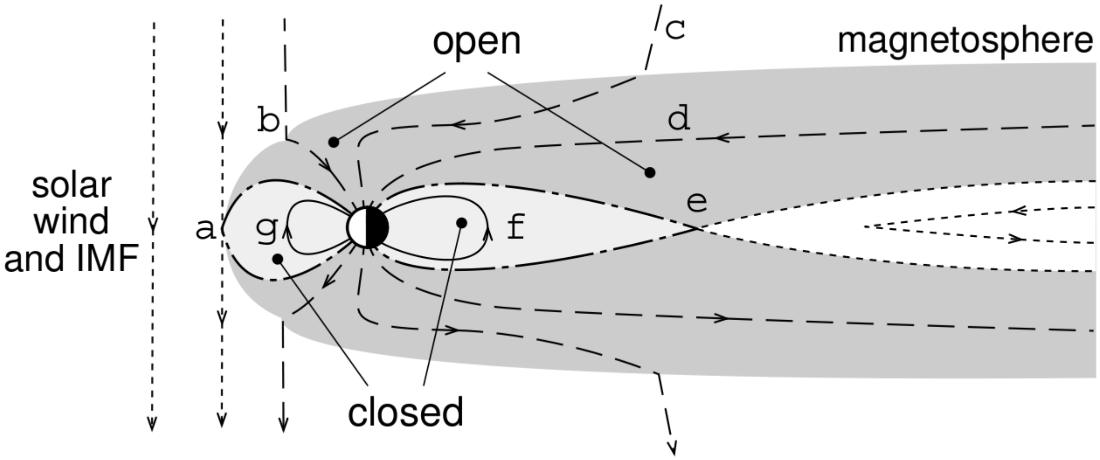


Figure 2.13: A diagram of magnetospheric convection during the Dungey cycle (Milan et al., 2003). IMF field lines are shown as dotted lines, dashed lines connect the Earth to the IMF and solid and dot-dashed lines represent the closed geomagnetic field.

The duration of a substorm is divided into three main phases; the growth phase, identified by McPherron (1970), followed by the expansion and recovery phases. The transport of magnetic flux that occurs during a substorm is known as magnetospheric convection or the *Dungey cycle*. Figure 2.13, from Milan et al. (2003), illustrates this process.

According to Rostoker et al. (1980), if the IMF is dominated by a northward B_z component for a prolonged period, the magnetosphere becomes quiet and approaches a state of lowest activity, which they term the *ground state*. Following a southward turning of the IMF, there is increased magnetic reconnection between the solar wind and magnetosphere (figure 2.13, *a*). This flux is transferred to the nightside (*b*, *c* and *d*), where a second stage of reconnection occurs to return to it the dayside (*e*). This energy transfer from the solar wind is seen in the atmosphere as an increase in polar cap magnetic flux and particle precipitation and causes the auroral convection region to expand equatorward. If the reconnection rate on the dayside exceeds that of the nightside then there is a build up of flux in the magnetotail, increasing its field strength (McPherron, 1979). This is termed a directly-driven process, due to its direct input from the solar wind.

Eventually the magnetotail is unable to support the excess energy, which marks the onset of the expansion phase. The point at which this occurs is associated with loading-unloading events, where energy stored within the magnetotail is transferred to the atmosphere. The directly-driven component of the substorm continues into this phase, which is responsible for the IMF-dependent high-latitude electric field in the ionosphere. The loading-unloading events, however, are driven by energy in the magnetotail and so are not related directly to the state of the IMF. These events are associated with a disruption in the cross-tail current, the exact cause of which is not universally agreed upon, however there are two leading theories. The near-Earth neutral line model (Baker et al., 1996) suggests that magnetic reconnection in the magnetotail accelerates particles within the plasma sheet toward and away from the Earth. The Earthward component is slowed by the increasing magnetic field, causing an eastward current and opposing the cross-tail current. The alternative theory is that of current-disruption (Lui, 1996), where near-Earth plasma instabilities disrupt the cross-tail current. This disruption can expand tailward and result in reconnection within the tail.

Finally the recovery phase occurs after the magnetosphere, and consequently the ionosphere, return to their pre-substorm structure. Aurora and plasma convection become weaker and contract poleward.

Storms and Substorms

A proxy typically used to characterise geomagnetic disturbances is the Disturbance Storm Time (Dst) index. The index is derived from magnetometer readings at low latitudes and measures the deviation of the horizontal component of the geomagnetic field. Whilst substorms typically affect the high-latitude region of the atmosphere, magnetospheric storms have a global impact and last for significantly longer. Storms are characterised by a main phase, followed by a gradual recovery. Table 2.1, from Gonzalez et al. (1994), shows how storms and substorms are dependent on the B_z component of the IMF.

	Dst (nT)	B_z **(nT)**	ΔT **(h)**
Intense	-100	-10	3
Moderate	-50	-5	2
Small (typical substorm)	-30	-3	1

Table 2.1: Characterisation of storm intensities for typical Dst and B_z values (from Gonzalez et al. (1994)). ΔT represents the time for which the B_z component must be sustained for an 80% chance of a given storm occurring.

A sudden and strong decrease in Dst, lasting several hours, occurs during the main phase, after which it gradually recovers to pre-storm levels over several days. For the most intense storms a large, prolonged, negative B_z is required, which may occur as the result of a large

Earth-impacting CME, or a strong CIR.

2.4.2 Coupling to the Neutral Atmosphere

The sources of energy input in to the Earth's atmosphere are from solar radiation (radiative heating) and charged solar wind particles (Joule heating and particle precipitation). These energy sources can increase dramatically as a result of solar variability, as seen in figure 2.14, due to events such as CIRs and CMEs. Radiative heating results from the various photo ionisation processes that are described at the start of this chapter. Particle precipitation is caused by charged particles penetrating in to the Earth's atmosphere and Joule heating is the result of ionospheric convection creating drag on the neutral atmosphere.

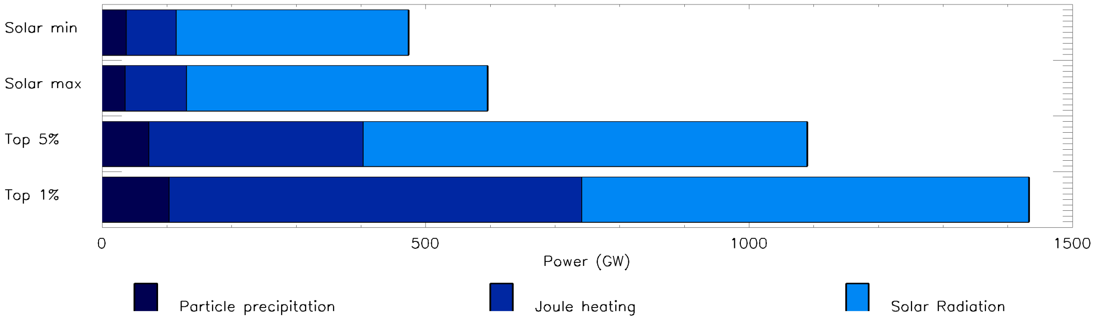


Figure 2.14: The relative contribution to the terrestrial energy input from solar radiation, Joule heating and particle precipitation (Knipp et al., 2005). The plots indicate solar minimum, solar maximum and the top 5% and 1% of space weather events in terms of power.

Radiative Heating

Absorption of solar radiation in the atmosphere is determined based on the Beer-Lambert law

$$I(\lambda) = I_\infty(\lambda)e^{-\tau(\lambda)} \tag{2.73}$$

where $I(\lambda)$ is the intensity of radiation at a wavelength, λ, and $I_\infty(\lambda)$ is the solar irradiance as a function of wavelength. $\tau(\lambda)$ is the optical depth, which determines the attenuation of solar radiation as it passes through the atmosphere. At an altitude of z_0 the optical depth is given by

$$\tau_{z_0}(\lambda) = \sum_j \sigma_j \int_{z_0}^{\infty} n_j(z) \sec(\chi) dz \tag{2.74}$$

where the sum is over different chemical species, j, $\sigma_j(\lambda)$ is the species and wavelength dependent absorption cross section, $n_j(z)$ is the height profile of the concentration of species j and the $sec(\chi)$ term accounts for radiation entering the atmosphere at an angle χ to the zenith. χ is dependent on season, latitude and local time, as given by

$$\cos(\chi) = \cos(\theta)\cos(\delta)\cos(H) + \sin(\theta)\sin(\delta) \tag{2.75}$$

where θ is the latitude, δ solar declination angle and H is the hour angle, 0° being local noon. This heating serves to create a large temperature gradient in the upper atmosphere, causing a circulation away from the dayside to the nightside, which is anti-sunward in the polar regions.

Joule Heating

A consequence of the magnetospheric processes described in section 2.4.1 is an ionospheric current system known as the *auroral electrojet*. These are caused by *Birkeland* currents (in fact, these are current sheets) that flow along geomagnetic field lines between the magnetosphere and ionospheric auroral regions. These currents flow both poleward, *region 1 currents*, and equatorward, *region 2 currents*, of the auroral oval in the dawn and dusk regions. On the dawn-side, region 1 currents flow into the ionosphere, whilst region 2 currents flow out of it. Conversely, on the dusk-side the flow is opposite. This is due to their connection to the oppositely directed Chapman-Ferraro and ring currents. This current closes across the auroral oval, whilst a small amount also closes over the pole. Transport of flux in the magnetosphere (section 2.4.1) therefore causes a convection of plasma that is mapped down to ionospheric altitudes in the auroral regions. This flow is anti-sunward over the pole and returns in a sunward direction at lower latitudes in the dawn and dusk regions. An ideal MHD plasma with a bulk flow velocity, $\mathbf{U}$, is described by the equation

$$\mathbf{E} = -\mathbf{U} \times \mathbf{B} \tag{2.76}$$

which means that the ionospheric convection is associated with an approximately horizontal electric field, an example of which is shown in figure 2.15 that has been produced using data from SuperDARN (section 3.2.1).

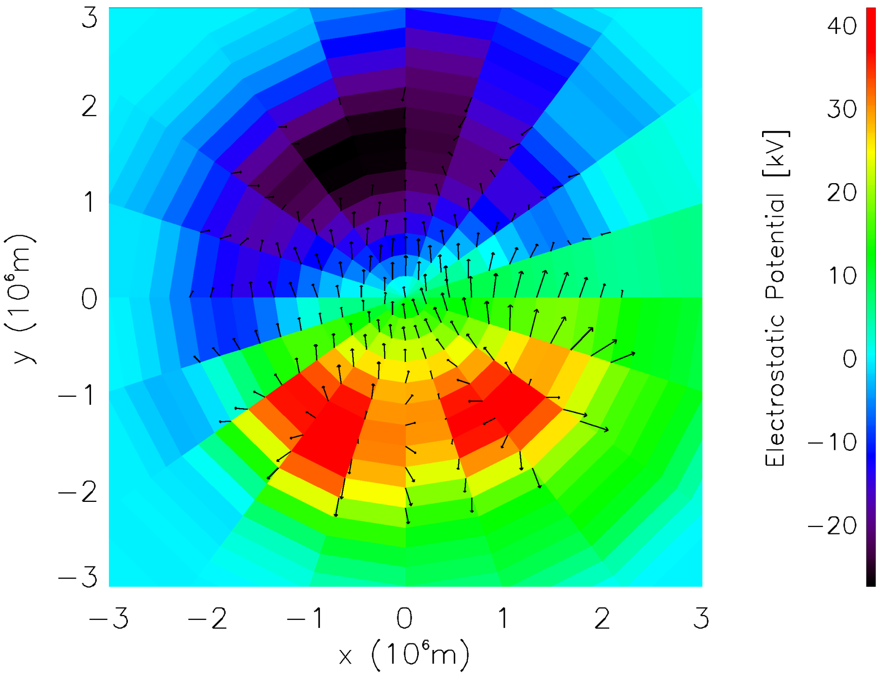


Figure 2.15: Northern polar ionospheric electric field, produced using data from SuperDARN, with 12MLT to the right of the plot. The charge distribution creates an electric field from dawn to dusk across the polar cap, resulting in an anti-sunward ion drift.

The Energy-Coupling Function

Despite the complex magnetospheric processes that connect the solar wind and ionosphere, Perreault and Akasofu (1978) were able to demonstrate a close correlation between IMF parameters and the energy generated by the solar wind-magnetosphere dynamo in terms of an empirical relationship called the *energy coupling function*,

$$\epsilon(t) = |\mathbf{v}||\mathbf{B}|\sin^4\left(\frac{\theta}{2}\right) r_{ms}^2 \tag{2.77}$$

where $\theta = \tan^{-1}(B_z/B_y)$ is the IMF clock angle. The radius of the magnetosphere, r_{ms}, in units of Earth radii, is assumed by the authors to have a constant value of $7R_e$. The strong dependence on the orientation of the magnetic field is because this is what governs the interaction between the IMF and the magnetosphere. Kan and Lee (1979) used this function to estimate the potential difference across the polar cap, Φ_{pc}, generated by the ionospheric electric field,

$$\Phi_{pc} = |\mathbf{v}||\mathbf{B}|\sin^2\left(\frac{\theta}{2}\right) r_{ms} \tag{2.78}$$

This has been developed over time to include more complicated physical processes. Work by Ridley (2005) has extended this formulation to include the density dependent effects of solar wind ram pressure on the position of the magnetopause, i.e. the size of r_{ms}. Russell et al. (2000) noticed a saturation in the polar ionospheric potential and Joule heating during the 24^{th} September 1998 geomagnetic storm. Siscoe et al. (2002) explained this as the effect of internal ionospheric currents, which act to reduce the strength of the subsolar magnetosphere and therefore inhibit reconnection. Ridley (2005) also included this in his formulation of the energy coupling function, in the form of the Alfvén Mach number, M_A,

$$\Phi_{pc} = \left(10^{-4}\mathbf{v}^2 + 11.7|\mathbf{B}|(1 - e^{\frac{M_A}{3}})\sin^3\left(\frac{\theta}{2}\right)\right)\frac{r_{ms}}{9} \tag{2.79}$$

$$M_A = |\mathbf{v}|\left(\frac{(\mu_0 n_e m_p)^{\frac{1}{2}}}{|\mathbf{B}|}\right) \tag{2.80}$$

$$r_{ms} = \left(\frac{(2B_s)^2}{2\mu_0 p}\right)^{\frac{1}{6}} \tag{2.81}$$

$$p = \frac{|\mathbf{B}|^2}{2\mu_0} + n_e m_p \mathbf{v}^2 \tag{2.82}$$

where n_e is the solar wind number density, μ_0 the permeability of free space, m_p proton mass, B_s the surface geomagnetic field strength and p the solar wind pressure. This formulation of the energy coupling function shows that energy transfer is dependent on solar wind speed, $\mathbf{v}$, density, ρ_e, and the IMF, $\mathbf{B}$.

For the purposes of general circulation modelling it is clearly important to quantify the state of the ionospheric electric field in a way that can be used to drive simulations. An early attempt by Foster et al. (1986) combined radar and satellite observations to quantify the geomagnetic activity into 10 levels of increasing power, each with a corresponding polar electric field and particle precipitation pattern. The electric fields were derived from ionospheric convection observed by the Millstone Hill radar, while the particle precipitation comes from observations by the TIROS and NOAA satellites (Fuller-Rowell and Evans, 1987). Weimer (1995) developed this idea by binning electric field observations from the Dynamics Explorer 2 satellite as a function of the IMF angle in the GSM[5] x-y plane. This was later extended to include the ionospheric response to variations in solar wind velocity (Weimer, 1996) and density (Weimer, 2001).

Particle Precipitation

Charged solar wind particles, mainly protons and electrons, that follow magnetic field lines can penetrate into the upper atmosphere at high latitudes. The particles transfer their kinetic energy directly when they collide with the neutral atmosphere. Figure 2.14 shows this contribution to the energy budget is less significant than Joule and radiative heating, however these particles have enough energy to cause ionisation and so contribute to the conductivity of the ionosphere, which in turn influences the amount of Joule heating. This energy deposition typically occurs between 100-200km at high latitudes and is responsible for the aurorae.

---

[5] Geocentric Solar Magnetospheric coordinates are centred on the Earth with the $x-$axis pointing toward the Sun. The $z-$axis lies at 90° in the plane containing the Earths magnetic axis, with magnetic north being positive. The $y-$axis completes the right handed coordinate system

2.4.3 Hemispheric Asymmetry in Ionosphere-Thermosphere Coupling

In addition to the dependence of the overall strength of the ionospheric electric field on the IMF B_z component, its state in both the northern (NH) and southern (SH) hemispheres is known to have a dependence on the B_y component. For $B_y > 0$ in the NH, the dusk cell (negative potential) increases in size and for $B_y < 0$, the dawn cell is increased. In the SH, the opposite is true; under $B_y < 0$ the dusk cell is enhanced and under $B_y > 0$ the dawn cell is enhanced (Heppner, 1972; Heppner and Maynard, 1987). This results in a potential difference between the NH and SH that has the same sign as B_y (Leontyev and Lyatsky, 1974) and which becomes greater when $|B_y| > |B_z|$ (Lu et al., 1994). The high conductivity of magnetospheric field lines means the closed lines short-circuit this potential difference, whilst regions of open field lines support the E_z field. This is the basis of models, such as Weimer (2005), which study potential distributions as a function of IMF sector. This is, however, an oversimplification of more complex processes and behaviour involving multiple cells can be observed during a transition between strong southward and northward IMF (Knipp et al., 1991). In addition, there is some delay between the IMF changing at the magnetopause and the changing convection. For example, Ridley et al. (1998) show that day-side convection can respond to changes in IMF at the magnetopause in a matter of seconds, whilst the response can be 10s of minutes on the night-side, along closed field lines (Khan and Cowley, 1999). Grocott and Milan (2014) study the response of ionospheric convection during periods where the IMF clock angle remains constant for a prolonged time. They find that averaged models of convection are a good representation when the IMF direction changes over periods of $\lesssim$ 30min, however, the observed convection deviates from these models when the same IMF conditions persist for longer. Whilst the IMF is generally a good proxy for the state of the ionospheric magnetic field at a given instant, this is not the case under many circumstances.

In addition to the effects of IMF B_y on the ionosphere, the geomagnetic field itself is asymmetric in each hemisphere in terms of both its strength and its offset from the geographic poles. The present day positions of the magnetic dip poles (from the IGRF (Finlay et al., 2010)) are shown in table 7.1, where the offset from the geographic pole is much larger in the SH. (Thayer and Killeen, 1993) used observations from the DE2 satellite to decompose the high latitude neutral winds into rotational and divergent components. The rotational component is shown to correspond to the ion drag and Coriolis forces, whilst the divergent component results from the pressure gradient caused by solar insolation and the combined effect of these two forces is dependent on the polar offset. One finding was that vortex formation in the neutral winds was consistently more dominant in the dusk sector, because of the divergent component points from ∼14-2MLT. Additionally, they found that the magnitude of neutral winds was generally larger in the SH where the geomagnetic field magnitude is greatest.

Chapter 3

Instrumentation and Models

The STEREO mission was conceived with the purpose of studying the causes and behaviour of CMEs and their propagation through the heliosphere. The mission consists of two spacecraft with identical instrumentation to measure the topology and physical properties of the solar wind between the Sun and the Earth. They follow heliocentric orbits with radii marginally inside (STEREO-A) and outside (STEREO-B) 1AU, causing them to drift relative to the Earth by approximately 22° per year.

With the exception of the two HELIOS spacecraft, past space-borne solar observatories have been positioned either in Earth orbit or at the Sun-Earth L1 point. These missions have aimed to follow CMEs through the heliosphere and to measure their effects in situ near the Earth. However, they suffer several limitations, particularly when applied to Earth-directed CMEs. When viewed from Earth, these CMEs appear as an expanding halo around the Sun, impeding measurements of their speed and direction, and therefore limiting arrival time predictions. In situ measurements from spacecraft near L1 provide detailed information about CMEs, but can only give about one hour's advance warning before they reach the Earth. The optically thin corona means that ambiguities about depth arise when trying to associate white light observations with density enhancements along a single line-of-sight. These limitations were the motivation for the STEREO mission and the three viewpoints provided by STEREO, plus Earth, allow them to be resolved.

3.1 The STEREO Mission

3.1.1 History

As knowledge of CMEs and their connection with geomagnetic storms developed throughout the first decade of observations in the 1970s, the solar physics community began to realise the benefits a multi-spacecraft mission would provide. The first such mission was proposed in 1982 (Schmidt and Bothmer, 1996), which would in fact have resembled the current STEREO mission quite closely. In 1992, two more missions were put forward, which would ultimately be merged to form the STEREO mission. The Solar Tomography Mission consisted of two or four spacecraft imaging the corona from different angles and preliminary simulations of tomographic reconstructions were published by Davila (1994). The Global Understanding of the Sun mission, described in Kaiser et al. (2008), proposed putting a single spacecraft in a heliocentric orbit, 90° from the Earth, with the aim of imaging Earth-directed CMEs. Due to the similarities between these missions, particularly their science objectives, they were united to create the STEREO mission (Davila et al., 1996). The initial concept involved four spacecraft: one at each of the L4 and L5 points, an Earth-orbiter and a fourth spacecraft in a heliocentric orbit inclined at 30° to the ecliptic. This was subsequently revised down to two spacecraft, resulting in the current STEREO mission.

3.1.2 Science Objectives

Four science goals were set out to fulfill the rather general primary objective of studying the origins and consequences of CMEs (Rust, 1999). In turn, several more specific measurement requirements were identified in order to address these goals (table 3.1).

Science Objective		**Measurement Requirement**
Understand the causes and mechanisms of CME initiation	A	Determine CME initiation time to within 10 minutes
	B	Determine location of initiation to within five degrees of solar latitude and longitude
Characterise the propagation of CMEs throughout the heliosphere	C	Determine the evolution of CME mass distribution and longitudinal extent to within five degrees as it propagates
	D	Determine the CME and MHD shock speeds to within 10% as it propagates
	E	Determine the direction of the CME and MHD shock propagation to within five degrees
Discover the mechanisms and site of energetic particle acceleration in the low corona and interplanetary medium	F	Develop distribution functions to an accuracy of 10% for electrons and/or ions with energies typical of solar energetic particles
	G	Locate regions of particle acceleration in the low corona to within 300,000km in radius and in interplanetary space to within 20 degrees in longitude
Develop a 3D time-dependent model of magnetic topology, temperature, density and velocity of the ambient solar wind	H	Obtain a time series of the solar wind temperature to within 10% accuracy at two points separated in solar longitude
	I	Obtain a time series of solar wind density to within 10% accuracy at two points separated in solar longitude
	J	Obtain a time series of solar wind speed to within 10% accuracy at two points separated in solar longitude
	K	Measure global magnetic field topology near the ecliptic by determining the magnetic field direction to within 10 degrees

Table 3.1: STEREO level 1 science requirements (table 1 from Kaiser et al. (2008)).

Both spacecraft achieving the measurements in table 3.1 for an interval of 150 days, after entering heliocentric orbit, was specified as the criterion for minimum success. Full success of the mission was regarded as the same measurements over a two year period, which was achieved on January 23, 2009.

3.1.3 Instrumentation

Four instrument packages were designed and developed in order to satisfy the measurement requirements, consisting of a total of 18 individual sensors (table 3.2). Each individual requirement is addressed by a combination of instruments from each spacecraft, so that, should a single instrument have failed, minimum success could still be achieved. The SECCHI instrument suite, which is responsible for remote sensing in white-light and UV, is the only one of the four that is used in this thesis and therefore the only one described here.

Instrument	**Acronym**	**Purpose**
SECCHI	COR1	Coronagraph 1.4-4.0 solar radii
	COR2	Coronagraph 2-15 solar radii
	EUVI	Extreme ultraviolet imager
	HI	Heliospheric imager 12-215 solar radii
IMPACT	SWEA	Solar wind electrons to 3 keV
	STE	Suprathermal electrons 2-100 keV
	SEPT	Electrons 20-400 keV; protons 60-7,000 keV
	SIT	Composition He-Fe 300-2,000 keV/nucleon
	LET	Protons, He, heavy ions to 40 MeV/nucleon
	HET	Protons, He to 100 MeV; electrons to 8 MeV
	MAG	Vector magnetic field to 65,536 nT
PLASTIC	SWS	Protons, alpha dist. functions to 100 keV
		Heavy ions to 100 keV
	WAP	Wide angle heavy ions to 100 keV
S/WAVES	HFR	Electric field 125 kHz-16 MHz
	LFR	Electric field 2.5-160 kHz
	FFR	Fixed frequency 32 or 34 M Hz
	TDS	Time domain to 250 k sample/sec

Table 3.2: STEREO instruments (table 2 from Kaiser et al. (2008)).

The five telescope Sun-Earth Connection Coronal and Heliospheric Investigation (SECCHI) instrument suite (Howard et al., 2008) is responsible for remote sensing observations in UV and visible light. The Extreme UltraViolet Imagers (EUVI) image the solar surface, whilst the visible light coronagraphs (COR-1 and COR-2) cover the low corona between $1.4R_\odot - 15R_\odot$. Each spacecraft is also equipped with two wide-angle, visible-light imagers to observe the heliosphere between the Sun and the Earth. HI-1 and HI-2 cover respective solar elongation angles of 4°-24° and 18.7°-88.7°, with their optical axes aligned to the ecliptic. This unique setup allows continuous, stereoscopic tracking of geoeffective CMEs in a field of view that contains the Earth. The HI design is based on the Solar Mass Ejection Imager (SMEI) (Eyles et al., 2003) on board the Coriolis spacecraft in Earth orbit.

	HI-1	**HI-2**
Direction of centre of field of view from Sun centre	14.0°	53.7°
Angular field of view	20°	70°
Angular range	$4° - 24°$ $(15R_\odot - 90R_\odot)$	$18.7° - 88.7°$ $(70R_\odot - 330R_\odot)$
CCD pixel size	35 arc-sec	2 arc-min
Image array (2×2 binning)	1024×1024	1024×1024
Image bin size	70 arc-sec	4 arc-min
Spectral bandpass	630-730 nm	400-1000 nm
Exposure time	40 seconds	50 seconds
Exposures per summed image sequence	30	99
Summed image cadence	40 minutes	2 hours
Brightness sensitivity ($B_\odot$ = solar disc)	$3 \times 10^{-15} B_\odot$	$3 \times 10^{-16} B_\odot$
Stray-light rejection (outer edge of field)	$3 \times 10^{-13} B_\odot$	$10^{-14} B_\odot$

Table 3.3: Performance specifications of the HI instruments (table 1 from Eyles et al. (2009)).

The major components of each HI assembly are a CCD detector, lens array, baffle system and a thermal radiator to cool the instrument below $-80°C$. The function of the baffle system is to protect the lenses from stray light due to the Sun and other out-of-field objects. This allows the instrument to operate with a brightness sensitivity adequate for observing diffuse density enhancements in the heliosphere. Details of the CCD detectors and HI fields of view are displayed in table 3.3 and detailed mechanical specifications of the instrument can be found in Eyles et al. (2009).

3.1.4 Mission Phases

Both STEREO-A and B were launched on October 26, 2006, into eccentric Earth orbits with several close flybys of the moon, making use of lunar gravity to achieve their respective heliocentric orbits (Driesman et al., 2008). For each of the instrument packages on board STEREO, there are differing angular spacings associated with optimum performance. For in situ particle and magnetometer measurements to be made from both spacecraft within a single magnetic cloud, a total separation angle of $< 50^\circ$ is required. Remote sensing observations in UV and white light work best when viewing features near the plane of the sky[1], suggesting a separation of $> 60^\circ$. The current orbit design, with each spacecraft continually separating, was selected as a compromise in order to satisfy all mission requirements. As a consequence the STEREO mission was divided into four phases (Kaiser et al., 2008). Phase 1 occurs whilst the total separation is $< 50^\circ$, during which time calibrations can be made between the instruments on both spacecraft. This phase is also optimal for viewing coronal structures. A separation of between 50° and 110° occurs in phase 2, which is optimal for triangulation of CMEs. This separation is also ideal for one spacecraft to observe a CME that impacts the other, allowing remote sensing measurements from the first spacecraft to be associated with in situ measurements from the second. Phases 3 and 4 occur after the main mission period, when the separation is beyond 110°. This configuration allows both spacecraft to observe Earth-directed CMEs in the plane of the sky. As the spacecraft separation passes 180° in phase 4, the entire solar surface is visible, allowing features on the far side of the Sun to be associated with activity on the near side.

3.1.5 Other Space-Based Instruments and Models

The OMNI Data Set

The High Resolution OMNI data consists of measurements from the ACE and WIND spacecraft, at the L1 position, and IMP-8 and Geotail in Earth orbit. ACE and WIND are in heliocentric orbits at an approximately constant distance of 1.5×10^6km from the Earth. Geotail and IMP-8 are both in eccentric Earth orbits with respective semi-major axes of $127,417$km and $218,146$km. The data are provided with a forward shifted time delay, to estimate the solar wind conditions at the bow shock of the magnetosphere, rather than the position of the observing spacecraft. The time shifting is performed using the solar wind

---

[1] The plane perpendicular to the observer's line-of-sight at any distance along it.

velocity vector and correcting for the offset of the observing spacecraft from the Sun-Earth line (Weimer et al., 2003) to provide estimates of IMF and solar wind parameters with 1 minute resolution.

WSA-ENLIL

ENLIL (Odstrcil and Pizzo, 1999; Odstrcil et al., 2004) is a three-dimensional magnetohydroynamic (MHD) model of the solar wind, which solves the time-dependent conservation equations of mass, momentum, energy density and of the magnetic field. The inner boundary lies between $21.5R_\odot$ and $30R_\odot$, which approximately coincides with the edge of the field-of-view of coronagraph instruments. The upper boundary may be adjusted, depending on the purpose of the simulation, and the angular extent is ±60° in heliospheric latitude, with 360° azimuthal coverage. The lower boundary is typically driven by the Wang-Sheeley-Arge (WSA) model. This is a combined empirical and physics-based model that supplies the ambient solar wind velocity, density, temperature and magnetic field. In order to model transient structures, such as CMEs, these are included as a perturbation to the background solar wind parameters at the lower boundary, which evolve though the heliosphere according to the model equations.

3.2 Atmospheric Observations

3.2.1 The Super Dual Auroral Radar Network

SuperDARN (Greenwald et al., 1995; Chisham et al., 2007) is an international network of High-Frequency (HF), coherent-scatter radars situated in both the northern and southern polar regions. The radars measure the line-of-sight component of ion velocity in the F-region of the ionosphere. Measurements from multiple radars are combined in order to create maps of polar ion convection with a time resolution of two minutes. If these velocities are assumed to result from an $\mathbf{E} \times \mathbf{B}$ drift, then the electrostatic potential pattern may be inferred via some assumptions about $\mathbf{E}$ and a known $\mathbf{B}$.

Radar Operation

Each SuperDARN radar operates by transmitting radio waves in the frequency range $8 - 20$MHz into the ionosphere's F-region (Ruohoniemi and Greenwald, 1996). The refractive index of the F-region varies spatially due to density irregularities. These irregularities

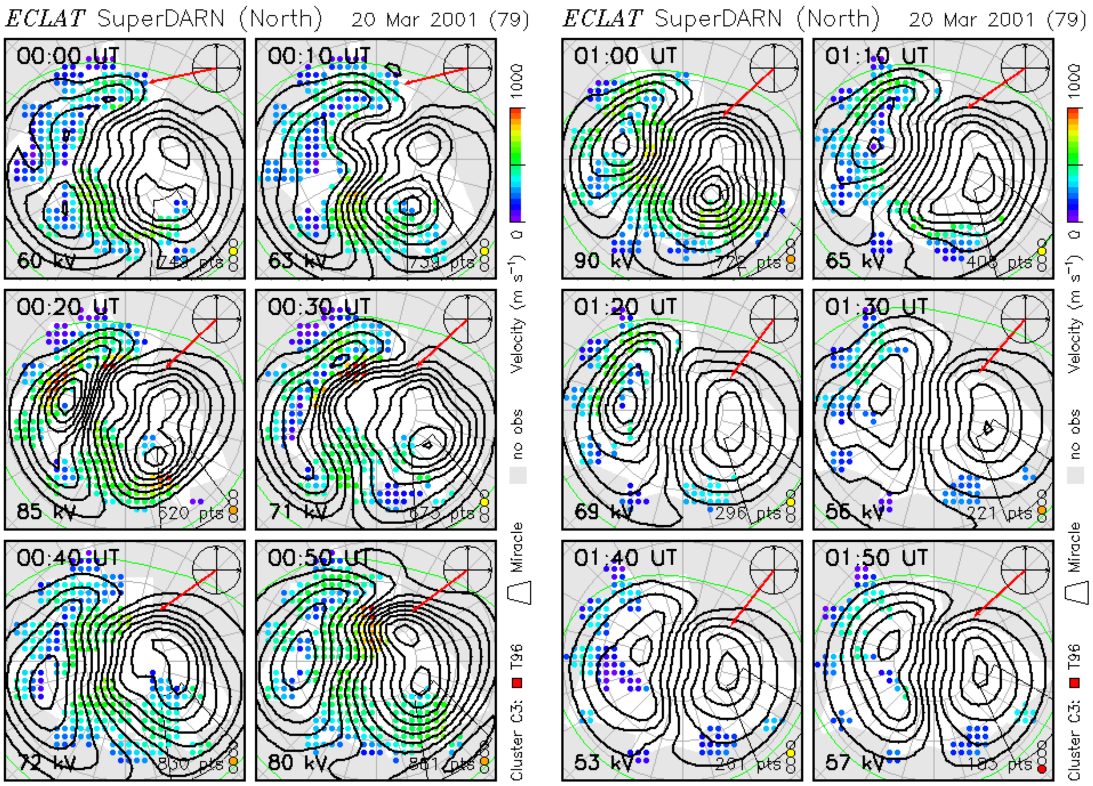


Figure 3.1: Example of SuperDARN potential maps, courtesy of the University of Leicester, from 20th March 2001. The coloured points indicate the contribution of radar observations to the resulting potential, which is represented by the contours. The red arrow indicates the IMF conditions, which are used to determine the statistical model used in the fitting procedure.

cause a back-scattered signal, which may then be detected by the radar. If the variations in refractive index scale as half the emitted wavelength, the back-scattered waves will constructively interfere. Bulk motion of ionospheric particles will then result in the back-scattered waves being Doppler shifted, allowing a measurement of line of sight speed. If two such radar beams overlap within a common volume they can be combined to produce a velocity vector. These vectors, along with the uncombined line of sight speeds, are then used to constrain a fitting procedure that produces a full potential pattern, expanded in terms of spherical harmonic functions. The fitting procedure used is that of Ruohoniemi and Greenwald (1996), which depends on the prevailing IMF conditions. Figure 3.1 shows an example of the contribution of the observations to the electrostatic potential distribution under particularly ideal conditions, when there is a large number of data points. The way the model changes over successive time-steps illustrates the basic difference from a model like Foster's. The final three panels (0130 to 0150UT) show how a lack of data causes the potential to resemble a more featureless shape, which begins to resemble Foster's model. The SuperDARN potential maps are produced using 49 spherical harmonic coefficients (6^{th} order) and the midnight latitude of the Heppner-Maynard Boundary (HMB). This boundary represents the countour at which electrostatic potential is equal to zero. Weimer's model is dependent on 39 spherical harmonic coefficients[2], whilst Foster's model is instead based on bin averaged velocities with a resolution of 0.5 hours local time by 2° magnetic apex latitude.

The angular coverage of each radar is divided into 16 sectors and the transmitted pulses are divided into 75 bins according to the distance at which scattering occurs. This results in a total of 1200 cells with a coverage of 52° in azimuth, a range of 3500km and a time resolution of 2 minutes. The data are available at three levels of processing. Level 1 data consist of the raw auto-correlation functions between transmitted and back-scattered signals, level 2 data result from a fitting technique to determine back-scattered power and Doppler shift and level 3 combines data from all radars to produce a map of electrostatic potential. It is the level 3 data that are used to drive CMAT2 in chapters 6, 7 and 8, which are made available courtesy of the Radio and Space Plasma Physics Group at the University of Leicester, for the years 2000-2010 every two-minutes in both hemispheres.

Electrostatic Potential in Terms of Spherical Harmonic Functions

Motion of ionospheric plasma is governed by the magnetospheric processes described in section 2.4. The associated electric field, **E**, is curl free and may therefore be described by the equation

$$\mathbf{E} = -\nabla\Phi \tag{3.1}$$

where Φ is an electrostatic potential, of which **E** is the negative of the gradient. The drift velocity may then be related to **E** and the geomagnetic field, **B**, by

$$\mathbf{U} = \frac{\mathbf{E} \times \mathbf{B}}{B^2} \tag{3.2}$$

The directions of the velocity vectors observed by SuperDARN therefore correspond to

---

[2] Weimer's model is calculated to 8^{th} order, but values of $|m| > 3$ are ignored (Weimer, 1995).

contours of equipotential. The electrostatic potential may be represented as an expansion in terms of spherical harmonic functions, Y_{lm}

$$\begin{aligned}\Phi(\theta',\phi) &= \sum_{l=0}^{L}\sum_{m=-l}^{l} A_{lm}Y_{lm}(\theta,\phi)\\ &= \sum_{l=0}^{L} A_{l0}P_l^0(\cos\theta) + \sum_{m=1}^{l}(A_{lm}\cos m\phi + B_{lm}\sin(m\phi))P_l^m(\cos\theta) \quad (3.3)\end{aligned}$$

where A_{lm} and B_{lm} are real-valued coefficients and θ' and ϕ are co-latitude[3] and magnetic local time, respectively. The spherical harmonic functions are expanded in terms of an angle, $0 \leq \theta \leq \pi$, however the potential is only defined above the HMB, Λ_0. The values of θ and θ' are therefore related by

$$\theta' = \frac{\pi}{\pi/2 - \Lambda_0}\cdot\theta \quad (3.4)$$

P_l^m are the associated Legendre polynomials given by:

$$P_l^m(x) = \frac{-1^m}{2^l l!}(1-x^2)^{m/2}\frac{d^{l+m}}{dx^{l+m}}(x^2-1)^l \quad (3.5)$$

which have known solutions for $P_l^m(\cos\theta)$ and are used in both the calculation of the electric and magnetic fields in CMAT2. Whilst the SuperDARN radars provide significant coverage of the high-latitude ionosphere, particularly in the northern hemisphere, during periods where back-scatter is reduced a means is required to account for the lack of data. An empirical model based on the three components of the IMF vector (B_x, B_y and B_z in GSM) is employed to constrain the spherical harmonic fit, in regions where there is no back-scatter (Ruohoniemi and Baker, 1998).

3.2.2 Incoherent Scatter Radars

Incoherent Scatter (IS) radars emit radio waves into the upper atmosphere and are able to measure some properties of the ionospheric plasma by observing the reflected signal. One

---

[3] The co-latitude at a point on the Earth's surface is the difference between π and the latitude, i.e. the angular distance from the north pole.

of these properties is electron density, n_e, which is determined by reflected power. The method described by Thayer (1998), which is summarised here briefly, may be used to infer further properties, such as conductivities and Joule heating. The Pedersen conductivity is given by (4.18) in section 4.1.1

$$\sigma_P = \frac{e n_e \nu_{in}}{B \Omega_i} \tag{3.6}$$

Where e is the electron charge, ν_{in} the ion-neutral collision frequency, B the magnetic field strength and Ω_i the ion cyclotron frequency. Taking ν_{in} from MSIS-90 and B from the IGRF, the radars are able to produce estimates of σ_P from measurements of n_e. Observing a range of altitudes permits the height integrated Joule heating to be calculated from

$$Q_J = \int_{z_{lower}}^{z_{upper}} \sigma_P(z) E^2 dz = \Sigma_P E^2 \tag{3.7}$$

where z is altitude and the upper and lower limits can be taken as the range of available data.

3.2.3 Fabry-Perot Interferometers

A Fabry-Perot Interferometer, or FPI, is designed to measure wavelengths at high spectral resolution in order to determine very small Doppler shifts of the order of 10^{-13}m in auroral and airglow emissions (e.g. Hecht et al. (1986)). The Fabry-Perot etalon uses two highly reflecting plates to allow multiple reflections and interference. The FPI etalon is placed on an optical bench with a telescopic system to focus the interferometer pattern onto a detector. This is used to measure airglow and auroral emissions. The most common wavelengths observed are the 630nm emission that has a peak emission altitude of around 240km, and the 557.7nm emission with a peak emission altitude around 120km. These allow direct measurements of the neutral winds and neutral temperatures from Doppler shifts and Doppler broadening of the emission spectral lines.

3.2.4 CHAMP

The Challenging Minisatellite Payload (CHAMP) was a atmospheric and ionospheric spacecraft that operated between 2000 and 2010 in a decaying geocentric orbit from approximately 450 to 300km. The payload included a tri-axial accelerometer, which could be used to infer neutral wind speeds (Liu et al., 2005). This is acheived using the following equation for satellite drag

$$\mathbf{a} = \frac{1}{2}\rho\frac{C_d}{m}AV^2\hat{\mathbf{v}} \tag{3.8}$$

Where $\mathbf{a}$ is accelaration from atmospheric drag, ρ, atmospheric density, C_d, the drag coefficient, A, effective area of the spacecraft, V relative velocity of the spacecrfat and $\hat{\mathbf{v}}$, a unit vector in the direction of V.

3.3 Numerical Atmospheric Models

The non-linear nature of the equations that describe the dynamics of the atmosphere cannot be solved analytically. Instead, it is possible to divide the atmosphere into discrete regions in each dimension over which a finite difference scheme may be used to solve the equations iteratively. The first such numerical model was devised as far back as 1913 by Lewis Fry Richardson (Lynch, 2008). However, without the means to solve such a large number of calculations, this ambition could not be realised at the time. It was not until 1950 that the first successful forecast was made. John von Neumann oversaw the design and construction of an electronic computer at Princeton and, familiar with the work of Richardson, turned it to the problem of weather forecasting.

Limited by computing power, the first General Circulation Models (GCMs) applied many simplifying assumptions to the physics of the atmosphere that were typically solved over a small region of the globe. Successive GCMs were expanded to include more complicated physics and chemistry and to cover larger areas of the atmosphere in three dimensions. These models are typically restricted to simulating either the lower or upper atmosphere, due to the very different processes that govern these regions. More recent *whole atmosphere* models have been developed by combining aspects of lower and upper atmosphere models, such as CMAT2. Due to the significant use of CMAT2 in this thesis, a thorough description is reserved for the following chapter. However, a summary of other relevant GCMs, and models that are used to drive them is given here.

3.3.1 General Circulation Models

GITM

The Global Ionosphere Thermosphere (GITM) model is a three-dimensional, phsyics-based GCM that simulated the ionosphere and thermosphere between 90 and 600km, with an adjustable resolution. Unlike CMAT2, and other thermospheric models, GITM uses an altitude-based vertical coordinate system, rather than a pressure-based one. This means the model is able to solve non-hydrostatic equations and therefore simulate vertical momentum transfer, allowing more realistic simulation of atmospheric dynamics. The model has been developed to use various electric field models, including that of Foster et al. (1986) used by CMAT2. However, GITM is typically run using the more advanced model of Weimer (2005).

TIE-GCM, LFM and CMIT

The design of the Thermosphere Ionosphere Electrodynamics General Circulation Model (TIE-GCM) bears many similarities to that of CMAT2; it solves mass, momentum and energy conservation equations over a three-dimensional, pressure-based grid, assuming hydrostatic equailibrium, constant gravity and incompressibility. The grid resolution is $5^\circ \times 5^\circ$ with 29 pressure-levels. Like CMAT2, the model uses the F10.7 index to determine solar EUV input and K_p to drive particle precipitation. However TIE-GCM differs in that it includes the IMF-dependent Weimer model in order to drive high-latitude electric fields. The Coupled Magnetosphere Ionosphere Thermosphere Model (CMIT) is a coupling of TIE-GCM to the Lyon-Fedder-Mobarry (LFM, Lyon et al. (2004)) magnetosphere model. The LFM model is a further three-dimensional MHD code that takes its input from the IMF vector and solar wind density and velocity.

3.3.2 High Latitude Electric Fields in GCMs

Measurements of the global ionospheric electric field date back to those of Heppner (1972) using NASA's OGO-6 satellite, which were used to produce the first empirical model of the high-latitude electric field (Heppner, 1977). An alternative model by Foster et al. (1986)

was based instead on ground based observations from the Millstone Hill radar, which were binned according to a precipitation index, equivalent to K_p. A third method by Kamide et al. (1981) used ground-based magnetometer readings, combined with assumptions about ionospheric conductivity, in order to infer the high-latitude electric field. The launch of the Dynamics Explorer mission in 1981 resulted in a more advanced satellite-based model by Heppner and Maynard (1987), and later in the IMF-dependent model of Weimer (1995). More recently, SuperDARN (Greenwald et al., 1995; Chisham et al., 2007) has greatly improved the coverage provided by ground-based observations.

The radar and satellite methods each have advantages and limitations, whilst the magnetometer approach has proven to be rather limited, due to it's dependence on other measurements. Satellite observations provide the only direct measurement of the ionospheric electric field, although they are restricted to point measurements along the orbital path. The empirical models created using satellite data therefore require a large number of orbits in order to build up a global picture of ion-convection. The weakness of this approach is that the models must be averaged and so contain no small scale spatial or temporal variability. This variability is important to modelling energy and momentum transfer to the thermosphere, a problem which is to be addressed in this chapter. The ground-based method, specifically SuperDARN observations, allows near-instantaneous coverage of the polar cap, in both hemispheres, with high time resolution. The weakness is that, in the absence of back-scattered signal, the network produces little data and becomes dependent on empirical models to fill in the gaps.

These measurements provide two potential ways of driving the high latitude electric field in GCMs; by means of a proxy-driven empirical model or by instantaneous measurements such as SuperDARN. A third method that may used is to couple an MHD model of the Earth's magnetosphere, such as the Grand Unified Magnetosphere Ionosphere Coupling Simulation (GUMICS) (Janhunen, 1996) or the Lyon-Fedder-Mobarry (LFM) model (Lyon et al., 2004), to the upper boundary of the GCM, however this does not currently exist in CMAT2.

Empirical models of the high latitude electric field are often used in general circulation modelling because they can be driven using a simple proxy, such as K_p or the IMF parameters. If these controlling parameters can be predicted it allows the model to be used as a forecasting tool. The unmodified version of CMAT2 uses Foster's model, whilst the more advanced Weimer model is also commonly used in other modelling studies. A brief description of each is given here.

Foster's Electric Field Model

An early attempt to create an empirical model of electric fields and particle precipitation by Foster et al. (1986) combined radar and satellite observations, respectively, to quantify the geomagnetic activity into 10 levels of increasing power, each with a corresponding polar electric field and particle precipitation pattern. The electric fields were derived from ionospheric convection observed by the Millstone Hill radar, whilst the particle precipitation comes from observations by the TIROS and NOAA satellites (Fuller-Rowell and Evans, 1987). Figure 3.2 displays the electrostatic potential distributions and electric fields from the six highest activity levels in Foster's model. This model is limited to a single hemisphere and therefore the same potential must be used in both hemispheres within CMAT2.

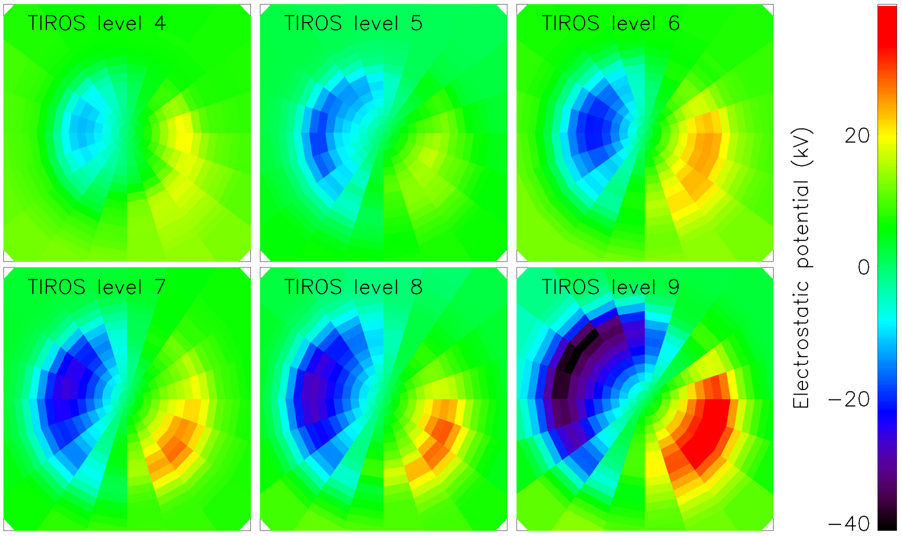


Figure 3.2: Electrostatic potential from Foster's empirical model, plotted over a grid in magnetic coordinates with 12 MLT to the top of the plot. The potential maps are shown for the six highest activity levels.

Weimer's Electric Field Model

Weimer (1995) developed Foster's idea further but instead binned electric field observations from the Dynamics Explorer 2 (DE2) satellite as a function of the IMF magnitude, $|\mathbf{B}|$, and orientation, $\theta = \tan^{-1}(B_z/B_y)$, in the GSM[4] x-y plane. This was later extended to include the ionospheric response to variations in solar wind speed, $|\mathbf{v}|$, (Weimer, 1996) and density, n_e, (Weimer, 2001). This provides a model based on four parameters, which are known, for example, from studies of the energy coupling function (Perreault and Akasofu, 1978; Kan and Lee, 1979; Akasofu and Ahn, 1980), to be closely linked to the ionospheric electric field.

3.3.3 Other Relevant Models

MSIS

The Mass-Spectrometer-Incoherent-Scatter (MSIS, Hedin et al. (1977a,b)) model is a description of the temperature and density of the upper atmosphere based on data from a number of sources, which was developed into the MSIS-90 model (Hedin, 1991) to extend the coverage from 0 to 700km altitude. CMAT2 uses the more recent NRLMSIS-00 model (Picone et al., 2002), which provides detailed representations of the neutral atmosphere through multiple data-sets including mass density from satellite accelerometers, temperature from IS radar and molecular composition from solar UV occultation as measured by the Solar Maximum Mission, which operated in low Earth orbit.

The K_p and $F10.7$ Indices

The K index takes a value ranging from 0 to 9 indicating increasing geomagnetic activity. It is based on the maximum fluctuations of magnetometer readings of the Earth's magnetic field during a three-hour period. The K_p index (Bartels et al., 1939) represents a weighted, planetary average of the K index based on a global network of magnetometer readings. Foster et al. (1986) show this index to correlate well with both particle precipitation power and cross-polar cap potential and as such it was often used as a proxy for both of these phenomena in older GCMs. The variation in solar microwave emissions at 10.7cm has been shown to correlate well both with sunspot number and with solar UV and visible power observed at Earth. Again, this commonly used as a proxy for energy input in GCMs, including CMAT2. The value of total solar irradianceat may be determined from F10.7, which is in turn used to calculate radiative heating in the atmosphere.

---

[4] Geocentric Solar Magnetospheric coordinates are centred on the Earth with the $x-$axis pointing toward the Sun. The $z-$axis lies at 90° in the plane containing the Earths magnetic axis, with magnetic north being positive. The $y-$axis completes the right handed coordinate system

The IGRF

The components of the magnetic field in the spherical eccentric dipole coordinate system can be found using

$$\mathbf{B} = -\nabla \left(\frac{-B_0 r_0^3}{r_m^2} \cos\theta_m \right) \tag{3.9}$$

where ∇ is the grad operator in spherical polar coordinates and B_0 is the reference magnetic field, determined from the dipole moment by

$$B_0 = \frac{\mu_0 M}{4\pi r_e^3} \tag{3.10}$$

The transformations required to convert the components of the field into the geographic frame are given in detail in Millward et al. (1996a). The geomagnetic field may be defined in terms of a magnetic scalar potential, γ. A spherical harmonic expansion of this can be written in the form

$$\gamma = r_0 \sum_{n=1}^{\infty} \sum_{m=0}^{n} \left(\frac{r_0}{r_g} \right)^{n+1} P_n^m(\cos\theta_g)(g_n^m \cos m\phi_g + h_n^m \sin m\phi_g) \tag{3.11}$$

where P_n^m are Schmidt normalised Legendre polynomials and g_n^m and h_n^m are the Gauss coefficients. The Gauss coefficients are calculated by fitting (3.10) and (3.11) to global observations from magnetic observatories. These have been calculated up to 14^{th} order for the years 1900-1995 and to 18^{th} order from 1995 onwards. This agreed upon set of Gauss coefficients is what makes up the IGRF.

This page was intentionally left blank

Chapter 4

The Coupled Middle Atmosphere Thermosphere Model (CMAT2)

The Coupled Middle Atmosphere-Thermosphere Model (CMAT2) general circulation model (GCM) is a physics based simulation of the dynamics of the Earth's atmosphere. The model is based upon fundamental equations of physics, and some simplifying assumptions, which are solved on a three dimensional global grid using a finite difference integration scheme. More precisely, the model solves the coupled, time-dependent equations of energy, momentum and mass conservation over a grid of points in longitude, latitude and pressure-defined altitude levels. The model solves for temperatures, velocities and composition of the atmosphere. This chapter contains details of the CMAT2 code, as it existed prior to any modifications made in later chapters. Much of the development of the CMAT2 within this thesis is concerned with the high-latitude electric fields models, and these changes are detailed in the relevant chapters; 6, 7 and 8.

CMAT2 combines various aspects of preceding GCMs, the first of which was a three dimensional time-dependent thermosphere model developed by Fuller-Rowell and Rees (1980) at UCL. This model had a lower boundary at 80km and solved for neutral wind and temperatures. This was then extended to include the composition of a two-species atmosphere, the light (O) and heavy (N_2 and O_2) major constituents (Fuller-Rowell and Rees, 1983) and was combined with the functionality of the Sheffield University Ionosphere model of Quegan et al. (1982) as described in Fuller-Rowell and Evans (1987). The model was then extended further to solve for the three major constituents individually (Fuller-Rowell, 1984), resulting in the Coupled Thermosphere-Ionosphere model (CTIM). CTIM was coupled with a mid and low latitude ionosphere/plasmasphere model by Millward et al. (1996a) to create the Coupled Thermosphere-Ionosphere-Plasmasphere model (CTIP). The first version of CMAT was a development of CTIM, where the lower boundary was extended down to 30km to include the mesosphere and with increased vertical resolution (Harris, 2001). CMAT2 is a re-written version of the original CMAT, whilst also including a lot of the functionality of CTIP as well.

4.1 Main Equations

The main equations within the model are based on three fundamental laws of physics; the conservation of mass, the conservation of momentum and the conservation of energy. In order to implement these equations in the model several assumptions must be made in order to simplify the physics; the model is assumed to be in a state of hydrostatic and local thermodynamic equilibrium, the atmosphere is assumed to be an ideal gas and the acceleration due to gravity is constant over the whole model grid.

4.1.1 Physics

Hydrostatic Equilibrium

A parcel of air is in hydrostatic equilibrium when the force of gravity it experiences is exactly balanced by the vertical pressure gradient acting on it. This generally is not true of the real atmosphere due to convective motion caused by variability in the distribution of heating and cooling throughout the atmosphere. Because the speed of these motions is very slow it can be assumed that the atmosphere is in a state of quasi-hydrostatic equilibrium,

$$\frac{\partial P}{\partial z} = -g\rho \tag{4.1}$$

where P is pressure, z is altitude, g is acceleration due to gravity and ρ is mass density.

Geopotential

Geopotential, Φ, is defined as the potential of the Earth's gravitational field, that is, the work done per unit mass in raising a parcel of air through a height dz,

$$d\Phi = gdz \tag{4.2}$$

The Ideal Gas Law

The equation of state for an ideal gas is

$$P = \frac{RT\rho}{M} = gH\rho \tag{4.3}$$

where R is the universal gas constant, T is the temperature of the gas, M is the mean molecular mass of the gas and H is the scale height, given by

$$H = \frac{RT}{Mg} \tag{4.4}$$

which is the step in altitude over which the pressure of the atmosphere decreases by a factor of e.

The First Law of Thermodynamics

The first law of thermodynamics is an alternative expression of the law of the conservation of energy. It states that the energy change of a system is equal to the net energy crossing the boundary of that system. For a compressible fluid this takes the form of

$$\frac{dT}{dt}C_p = \frac{1}{\rho}\frac{dP}{dt} + Q \tag{4.5}$$

where t is time and C_p is the specific heat capacity at constant pressure. The left hand term is the change in internal energy of the system, the first right hand term is the work done by the system and Q is the change in heat due to all sources and sinks.

The Continuity Equation

The continuity equation results from the law of the conservation of mass. For an infinitesimal volume element $\Delta x \Delta y \Delta z$, the mass inflow along the x axis is $\rho v_x \Delta y \Delta z$. The outflow from the opposite face is then $(\rho v_x + \Delta x \frac{\partial \rho v_x}{\partial x})\Delta y \Delta z$. The rate of change of mass per unit volume is then given by the difference, $-\frac{\partial \rho v_x}{\partial x}$. The rate of change of mass per unit volume in 3 dimensions is therefore given by

$$\frac{\partial \rho}{\partial t} = -\nabla_3 \cdot \mathbf{V}_3 \rho \tag{4.6}$$

where ∇_3 is the three dimensional del operator and $\mathbf{V}_3$ is the three dimensional velocity vector in the (x, y, z) coordinate system.

The Momentum Equation

The equation of the conservation of momentum follows from Newton's second law, as applied to the neutral atmosphere by Rishbeth and Garriott (1969)

$$\frac{d\mathbf{V}}{dt} = -2\mathbf{\Omega} \times \mathbf{V} + \mathbf{g} - \frac{1}{\rho}\nabla P + \frac{\mu}{\rho}\nabla(\nabla \cdot \mathbf{V}) - \nu_{ni}(\mathbf{V} - \mathbf{U}) \tag{4.7}$$

where the right hand terms are the Coriolis force, the acceleration due to gravity, the pressure gradient,the viscous drag and the ion-neutral drag. $\mathbf{\Omega}$ is the Earth's angular rotation vector, $\mathbf{g}$ the acceleration due to gravity, μ the sum of the molecular and turbulent coefficients of viscosity, ν_{ni} the neutral-ion collision frequency and $\mathbf{U}$ the ion velocity vector. Although the above equation is three dimensional, the vertical component is small enough that it can be ignored. The vectors $\mathbf{V}_3$, $\mathbf{U}_3$ and ∇_3 are replaced by their two dimensional equivalents $\mathbf{V}$, $\mathbf{U}$ and ∇.

The Coriolis effect is the cause of the apparent deflection of moving objects when viewed from a rotating frame of reference. This can be thought of as a fictitious force that does no work but acts to alter the direction of the object with respect to the rotating frame. The force acts in a direction perpendicular to both the rotation axis and the direction of motion of the object. In the case of the model the Earth's spin provides a rotating reference frame and so air moving within the atmosphere experiences a Coriolis force. This force produces clockwise motion in the northern hemisphere and anticlockwise motion in the southern hemisphere.

The horizontal component of the Coriolis force is given by $f\mathbf{V} \times \hat{\mathbf{k}}$ where the Coriolis parameter, f, is equal to $2\Omega_T \sin\theta$, $\hat{\mathbf{k}}$ is a unit vector in the direction of the Earth's axis of rotation and θ is latitude. Ω_T is the magnitude of the total angular rotation rate, which is the Earth's rotation rate, Ω, plus a contribution from the fact that motion within the atmosphere is restricted to movement on a spherical surface. This is given by

$$\Omega_T = 2\Omega + \frac{V_y}{r_0 \cos\theta} \tag{4.8}$$

where V_y represents movement in an eastward direction and r_0 is the Earth's radius.

The second right hand term in (4.7) is the acceleration due to the force of gravity. This is kept constant throughout the whole model grid with a value of 9.5ms^{-2}, equivalent to an altitude of 130km.

The third right hand term of (4.7) represents the force resulting from pressure gradients. These are produced by heating and cooling in the atmosphere and are the main force driving neutral winds. The force will act in the opposite direction to the pressure gradient so that air will move from regions of high to low pressure.

The fourth term in (4.7) accounts for atmospheric viscosity. This force serves to smooth out any velocity gradients or *wind shears.* Wind shears in the atmosphere are typically greater in the vertical direction than the horizontal and so it is the vertical component of this term that dominates.

The final term in (4.7) is the drag force between the neutral and ionised components of the atmosphere. This can cause both acceleration and deceleration of the neutrals due to the driving forces governing the behaviour of the ionosphere.

The ion drag term can be obtained by starting with the single species momentum equation taken from Rishbeth (1972)

$$\frac{d\mathbf{U}_3}{dt} = \mathbf{g} + \nabla\left(\frac{P_i + P_e}{nm_i}\right) + \frac{e}{m_i}(\mathbf{E} + \mathbf{U}_3 \times \mathbf{B}) - \nu_{in}(\mathbf{U}_3 - \mathbf{V}_3) \tag{4.9}$$

where P_i and P_e are the ion and electron pressures, $n = n_i = n_e$ the electron/ion number density, m_i the ion mass, e the electron charge, $\mathbf{E}$ the electric field vector, $\mathbf{B}$ the magnetic field vector and ν_{in} the ion-neutral collision frequency. Here, the assumptions $m_i \gg m_e$ and $\nu_{in} = \nu_{en}$ have been made. The Coriolis term has been neglected from this form of the momentum equation because $\Omega \ll \nu_{in}$. We can ignore the acceleration term by averaging over many gyrocycles and, assuming steady state and taking just the horizontal components of (4.9), we can write

$$en_i(\mathbf{E} + \mathbf{U} \times \mathbf{B}) - \rho_i \nu_{in}(\mathbf{U} - \mathbf{V}) = 0 \tag{4.10}$$

$$-en_e(\mathbf{E} + \mathbf{U}_e \times \mathbf{B}) - \rho_e \nu_{en}(\mathbf{U}_e - \mathbf{V}) = 0 \tag{4.11}$$

where $\rho_i = nm_i$ and $\rho_e = nm_e$ are the ion and electron mass densities and $\mathbf{U}_e$ the electron velocity. Assuming charge neutrality, $en_i = en_e$, (4.10) and (4.11) become

$$en_i(\mathbf{U} \times \mathbf{B}) - en_e(\mathbf{U}_e \times \mathbf{B}) - \rho_i \nu_{in}(\mathbf{U} - \mathbf{V}) - \rho_e \nu_{en}(\mathbf{U}_e - \mathbf{V}) = 0 \tag{4.12}$$

The current density, $\mathbf{J}$, is defined as electric current per unit area, which, in a single species plasma can be written

$$\mathbf{J} = en_i\mathbf{U} - en_e\mathbf{U}_e \tag{4.13}$$

This can be used to rewrite (4.12) in terms of current density as

$$\mathbf{J} \times \mathbf{B} - \rho_i \nu_{in}(\mathbf{U} - \mathbf{V}) - \rho_e \nu_{en}(\mathbf{U}_e - \mathbf{V}) = 0 \tag{4.14}$$

from which we arrive at the expression for ion drag in the momentum equation,

$$\nu_{in}(\mathbf{U} - \mathbf{V}) = \frac{1}{\rho_i}\mathbf{J} \times \mathbf{B} - \frac{\rho_e}{\rho_i}\nu_{en}(\mathbf{U}_e - \mathbf{V}) = 0 \tag{4.15}$$

Neglecting the second right term from the assumption $\rho_i \gg \rho_e$ this becomes

$$\nu_{in}(\mathbf{U}-\mathbf{V}) = \frac{1}{\rho_i}\mathbf{J}\times\mathbf{B} \tag{4.16}$$

Current density in the presence of both electric and magnetic fields is given by a generalisation of Ohm's law

$$\mathbf{J} = \sigma\cdot(\mathbf{E}+\mathbf{U}\times\mathbf{B}) \tag{4.17}$$

E and **B** are found from the electric and magnetic field models described in section 4.2 and σ is the 3×3 conductivity tensor. This can be reduced to a 2×2 tensor (Rishbeth and Garriott, 1969), which is a function of magnetic dip angle (the angle a magnetic field line makes to the horizontal), I, and σ_0, σ_1 and σ_2, the altitude dependent parallel, Pedersen and Hall conductivities. σ_0 is the conductivity for currents in the direction parallel to **B** and is the dominant term above about 400km. Pedersen conductivity, σ_1, is the conductivity relating to currents in the direction parallel to **E** but perpendicular **B** and is dominant at around 130km. The Hall conductivity, σ_2, is perpendicular to both **E** and **B** and peaks at around 100km. The values of σ_1 and σ_2 are determined from

$$\sigma_1 = \frac{n_i e r}{B(1+r)} \tag{4.18}$$

$$\sigma_2 = \sigma_1 r \tag{4.19}$$

where

$$r = \frac{m_i \nu_{in}}{eB} \tag{4.20}$$

is the ratio of the ion-neutral collision frequency to the ion-gyrofrequency. The values of collision frequencies for the molecular ions, NO^+ and O_2^+ and taken from Schunk and

Walker (1973), while for O^+ the collision frequency is from Salah (1993),

$$\nu_{in-molecular} = 4.34 \times 10^{-6}[N_2] + 4.28 \times 10^{-16}[O_2] + 2.44 \times 10^{-6}[O] \tag{4.21}$$

$$\begin{aligned}\nu_{in-O^2} &= 6.82 \times 10^{-16}[N_2] + 6.66 \times 10^{-16}[O_2] \\ &+ 3.42 \times 10^{-17} T^{0.5}(1.08 - 0.139 log_{10}T + 4.51 \times 10^{-3}(log_{10}T)^2)[O]\end{aligned} \tag{4.22}$$

The 2×2 conductivity tensor is given by

$$\sigma_l = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{pmatrix} \tag{4.23}$$

where the components are given by

$$\sigma_{xx} = \frac{\sigma_0 \sigma_1}{\sigma_1 \cos^2 I + \sigma_0 \sin^2 I} \approx \frac{\sigma_1}{\sin I} \tag{4.24}$$

$$\sigma_{xy} = -\sigma_{yx} = \frac{\sigma_0 \sigma_2 sinI}{\sigma_1 \cos^2 I + \sigma_0 \sin^2 I} \approx \frac{\sigma_2}{\sin I} \tag{4.25}$$

$$\sigma_{yy} = \frac{\sigma_0 \sigma_1 \sin^2 I + (\sigma_1^2 + \sigma_2^2)\cos^2 I}{\sigma_1 \cos^2 I + \sigma_0 \sin^2 I} \approx \sigma_1 \tag{4.26}$$

The simplifying approximations can be made because I is close to 90°.

The Energy Equation

The first law of thermodynamics represents the law of the conservation of energy and is expressed in (4.5). This says that the change of internal energy of a system must equal the net amount of energy crossing the boundary of that system. In order to apply this to the atmosphere it is necessary to arrive at an equation that accounts for all of the sources and sinks of energy within the atmosphere, all of the means by which energy is transported and any processes that convert between different types of energy. An expression for kinetic energy arises by multiplying the momentum equation by **V**, since

$$\mathbf{V} \cdot \frac{d\mathbf{V}}{dt} = \frac{1}{2}\frac{d\mathbf{V}^2}{dt} \tag{4.27}$$

The energy equation is then found by summing kinetic energy, internal energy and geopotential. This is expanded on in section 4.1.3, which shows the energy equation in the form that is used by CMAT2.

4.1.2 Coordinate System

The numerical computation is performed over a grid of longitude, latitude and pressure-defined altitude levels. Because the equations in the previous section have been expressed in Cartesian coordinates it is necessary to now express them in terms of the new pressure based coordinate system. The reason we do this is because the assumption of hydrostatic equilibrium allows us to simplify the mathematics somewhat, which should become apparent in the following section.

The Cartesian coordinate system used thus far is defined as x is positive southward, y positive eastward and z positive upward. In the new spherical pressure based coordinate system latitude, θ, is positive southward, longitude, ϕ, is positive eastward and the pressure coordinate, p, is positive upward. The default model grid comprises 20 cells zonally (18° each), 91 cells meridionally (2° each) and 63 cells vertically, defined by

$$P = P_0 e^{\frac{-z}{H}} = P_0 e^{\Delta_n(1-n)} \tag{4.28}$$

where P_0 is the pressure of the bottom level, n an integer from 1 to 63 and Δ_n has a value of $\frac{1}{3}$ for levels 1 to 60 and a value of 1 for the top three levels. The model grid is fixed relative to the Earth and is therefore a rotating non-inertial reference frame.

The Lagrangian derivative

When considering moving fluids it is important to make the distinction between *Eulerian* and *Lagrangian* coordinate systems. In a Lagrangian frame of reference the observer moves with the fluid, that is, the observer follows a fluid parcel as its position in space changes. In an Eulerian frame of reference the observer focuses on a location through which the fluid passes. In the model the grid is fixed to the surface of the Earth, which the atmosphere flows relative to. Our coordinate system is therefore Eulerian. The derivative in the left hand term of the momentum equation (4.7) is the Lagrangian derivative, which must be converted to Eulerian form using

$$\frac{dA}{dt} = \frac{\partial A}{\partial t} + (\mathbf{V}_3 \cdot \nabla_3)A \tag{4.29}$$

The first right hand term is the derivative in the Eulerian frame, while the second is the advective term that represents the carrying properties of the fluid.

Horizontal Derivative

The transformation from z to p coordinates in the x direction for an arbitrary quantity, S, is given by

$$\left(\frac{dS}{dx}\right)_p = \left(\frac{dS}{dx}\right)_z + \left(\frac{dz}{dx}\right)_p \left(\frac{dS}{dz}\right) \tag{4.30}$$

Using the expressions for geopotential (4.2) and hydrostatic equilibrium (4.1) we arrive at

$$\left(\frac{\partial S}{\partial x}\right)_z = \left(\frac{\partial S}{\partial x}\right)_p + \rho \left(\frac{\partial \Phi}{\partial x}\right)_p \left(\frac{\partial S}{\partial P}\right) \tag{4.31}$$

which gives the gradient of S in the x direction in the new coordinate system. The same applies to the y direction, so the two dimensional del operator becomes

$$\nabla_z \cdot \mathbf{A} = \nabla_p \cdot \mathbf{A} + \rho \nabla_p \Phi \cdot \frac{\partial \mathbf{A}}{\partial P} \tag{4.32}$$

Vertical Derivative

Similarly, in the z direction we can again use (4.1) to find the vertical gradient of S

$$\frac{\partial S}{\partial z} = \frac{\partial S}{\partial P}\frac{\partial P}{\partial z} = -g\rho\frac{\partial S}{\partial P} \tag{4.33}$$

Velocity

Vertical velocity in the new coordinate system is given by

$$w = \frac{dP}{dt} \tag{4.34}$$

where w is positive in a downward direction relative to an isobaric surface. Using the expression for the Lagrangian derivative (4.29), $\frac{d}{dt}$ becomes

$$\frac{d}{dt} = \left(\frac{\partial}{\partial t}\right)_p + (\mathbf{V}_p \cdot \nabla_p) + w\frac{\partial}{\partial p} \tag{4.35}$$

where the del operator has been separated into horizontal and vertical components. When applied to geopotential (4.2) this becomes

$$\frac{d\Phi}{dt} = g\frac{dz}{dt} = \left(\frac{\partial \Phi}{\partial t}\right)_p + \mathbf{V}_p \cdot \nabla_p \Phi + w\frac{\partial \Phi}{\partial p} \tag{4.36}$$

That is,

$$\frac{dz}{dt} = \frac{1}{g}\left(\frac{\partial \Phi}{\partial t}\right)_p + \frac{1}{g}(\mathbf{V}_p \cdot \nabla_p \Phi) + \frac{w}{g}\frac{\partial \Phi}{\partial p} \tag{4.37}$$

where $\frac{dz}{dt}$ is the vertical velocity, v_z, in the Cartesian coordinate system. The first two right hand terms are an expression of *barometric vertical velocity*, which is the motion of an isobaric surface in the z direction due to expansion and contraction. The last right hand term is the *divergence vertical velocity* or the vertical velocity relative to an isobaric layer due to convergences or divergences in horizontal velocity.

4.1.3 Transformed Equations

Using (4.32) and (4.2) to express horizontal and vertical derivatives in the new coordinate system allows the fundamental equations in section 4.1 to be transformed. In some cases it is also necessary to convert from the Lagrangian to Eulerian frame of reference using (4.29).

Hydrostatic Equilibrium

Using (4.33) and (4.1), the equation of hydrostatic equilibrium becomes

$$\frac{\partial \Phi}{\partial P} = -\frac{1}{\rho} \tag{4.38}$$

The Ideal Gas Law

The equation of state for an ideal gas is already expressed in terms of pressure and so (4.3) remains unchanged

$$P = \frac{RT\rho}{M} = gH\rho \tag{4.39}$$

The Continuity Equation

Using (4.32) and (4.33) the continuity equation (4.6) becomes

$$\frac{\partial w}{\partial p} + \nabla_p \cdot \mathbf{V}_p = 0 \tag{4.40}$$

The Momentum Equation

The momentum equation (4.7) in the new coordinate system becomes

$$\begin{aligned} \frac{\partial \mathbf{V}_p}{\partial t} = -\mathbf{V}_p \cdot \nabla_p \mathbf{V}_p - w\left(\frac{\partial \mathbf{V}_p}{\partial P}\right) - \nabla_p \Phi - \left(2\Omega + \frac{V_y}{R\sin\theta}\right)\cos\theta \hat{\mathbf{k}} \times \mathbf{V}_p \\ + g\frac{\partial}{\partial P}\left((\mu_m + \mu_t)\frac{P}{H}\frac{\partial \mathbf{V}_p}{\partial P}\right) + \frac{\mu_m}{\rho}\nabla_p^2 \mathbf{V}_p + \frac{1}{\rho}\mathbf{J} \times \mathbf{B} \end{aligned} \tag{4.41}$$

where μ_m and μ_t are the molecular and turbulent coefficients of viscosity respectively, $\mathbf{J}$

the current density and $\mathbf{B}$ the Earth's magnetic field. When split in to the zonal, V_ϕ, and meridional, V_θ, components in the spherical coordinate system this becomes

$$\begin{aligned}\frac{\partial V_\phi}{\partial t} = & -\frac{V_\theta}{R}\frac{\partial V_\phi}{\partial \theta} - \frac{V_\phi}{R\sin\theta}\frac{\partial V_\phi}{\partial \phi} - \omega\frac{\partial V_\phi}{\partial p} - \frac{g}{R\sin\theta}\frac{\partial h}{\partial \theta} \\ & - \left(2\Omega + \frac{V_\phi}{R\sin\theta}\right) V_\theta cos\theta + g\frac{\partial}{\partial p}\left((\mu_m + \mu_t)\frac{p}{H}\frac{\partial V_\phi}{\partial p}\right) - \nu_{ni}(V_\phi - U_\phi) \qquad (4.42)\end{aligned}$$

$$\begin{aligned}\frac{\partial V_\theta}{\partial t} = & -\frac{V_\theta}{R}\frac{\partial V_\theta}{\partial \theta} - \frac{V_\phi}{R\sin\theta}\frac{\partial V_\theta}{\partial \phi} - \omega\frac{\partial V_\theta}{\partial p} - \frac{g}{R}\frac{\partial h}{\partial \theta} \\ & - \left(2\Omega + \frac{V_\phi}{R\sin\theta}\right) V_\phi cos\theta + g\frac{\partial}{\partial p}\left((\mu_m + \mu_t)\frac{p}{H}\frac{\partial V_\theta}{\partial p}\right) - \nu_{ni}(V_\theta - U_\theta) \qquad (4.43)\end{aligned}$$

A full derivation of (4.41), (4.42) and (4.43) is presented in Fuller-Rowell and Rees (1980) and Fuller-Rowell (1981).

The Energy Equation

Expressions relating the energy density of a gas, the potential energy per unit mass and the rate of change of energy are used to derive a term for the energy equation (4.27) in the new coordinate system

$$\begin{aligned}&\frac{\partial \epsilon}{\partial t} + \mathbf{V}_p \cdot \nabla_p(\epsilon + gh) + w\frac{\partial(\epsilon + gh)}{\partial p} = Q_{SOL} + Q_{IR} + \\ &g\frac{\partial}{\partial p}\frac{(K_M + K_T)}{H}p\frac{\partial T}{\partial p} + \frac{1}{\rho}(K_M + K_T)\nabla_p^2 T - g\frac{\partial}{\partial p}\frac{K_T g}{C_p} + \\ &\qquad \frac{\mathbf{J}\cdot\mathbf{E}}{\rho} + g\frac{\partial}{\partial p}\frac{\mu}{H}p\mathbf{V}\cdot\frac{\partial \mathbf{V}}{\partial p} \qquad (4.44)\end{aligned}$$

where $\epsilon = \frac{1}{2}(\mathbf{V}_p^2) + C_p T$ is the kinetic energy plus the specific enthalpy per unit mass, Q_{SOL} solar heating, Q_{IR} infra red cooling and K_M and K_T are the coefficients of molecular and turbulent conductivity.

4.1.4 Integration Technique

The non-linear nature of the equations presented in section 4.1.3 means they cannot be solved analytically. Instead, the model employs a finite difference scheme. That is, the model consists of a three dimensional grid of latitude, longitude and pressure-defined altitude levels upon which the equations can be solved at each point.

Finite Difference Method

The equations of continuity, momentum and energy are mapped on to the model grid so that

$$F_j = F(x_j) \tag{4.45}$$

where F_j is that value of the function F at grid point x_j, defined by

$$x_j = x_1 + \sum_{n=1}^{j-1} \Delta x_n \tag{4.46}$$

for the domain $x_1 \leq x_j \leq x_N$, where N is the number of grid points. Using this method the first order spatial derivative can be approximated in each dimension by taking the central difference of the two neighbouring cells in that dimension

$$\Delta'_x F_j = \frac{dF_j}{dx} = \frac{F_{j+1} - F_{j-1}}{2\Delta x_j} \tag{4.47}$$

where Δx_j represents the grid step length. The approximated second order derivative is then given by

$$\Delta''_x F_j = \frac{d^2 F_j}{dx^2} = \frac{F_{j+1} - 2F_j + F_{j-1}}{(\Delta x_j)^2} \tag{4.48}$$

The equations of momentum and energy are of the form

$$\frac{\partial F(t)}{\partial t} = LF(t) \tag{4.49}$$

where L represents the operator containing first and second order derivatives. At a time step t, the model parameters are determined from the time step $t-1$ by integrating the physics equations using (4.47) and (4.48). This iterative process is then carried out over a finite time step Δt, which allows (4.49) to be solved using

$$F_{n+1} = F_n + LF_n\Delta t \tag{4.50}$$

where the subscript n indexes the time step.

Numerical Instabilities

The approximation method described here is known as the Euler method. It is unstable for first order differential equations of the form

$$\frac{\partial F}{\partial t} + v\frac{\partial F}{\partial x} = 0 \tag{4.51}$$

such as advective terms, which appear in the momentum and energy equations. The method is, however, stable for second order differential equations, such as diffusive terms of the form

$$\frac{\partial F}{\partial t} + K\frac{\partial^2 F}{\partial x^2} = 0 \tag{4.52}$$

where K is a diffusion coefficient. This stability holds so long as the time step satisfies the condition

$$\Delta t < \frac{1}{2K}(\Delta x)^2 \tag{4.53}$$

The instability in first order derivatives is prevented by introducing a numerical diffusion term to (4.51) so that it takes the form

$$F_{n+1} = F_n + LF_n\Delta t = \frac{1}{2}\Delta t(\mathbf{V}\cdot\nabla)(\mathbf{V}\cdot\nabla)F_n \tag{4.54}$$

This modified Euler scheme is known as an explicit forward time central space scheme and is stable for equations containing both first and second order derivatives.

Towards the upper boundary of the model molecular diffusion becomes the dominant transport process and the time step Δt required to satisfy the condition (4.53) becomes very small. To avoid the computational costs of reducing the time step, instead the scale height is increased. The vertical resolution of the top three levels of the model grid are increased from 1/3 to 1 scale height.

4.1.5 Boundary Conditions

Upper Boundary

The temperatures and velocities at the very top level of the model are set equal to the level below. It is assumed that there are no momentum or energy sources at the upper boundary and therefore the temperature and velocity gradients will become zero. It is also assumed that no mass can flow vertically out of the top of the model, so the vertical velocity, w, at the upper boundary is set to zero.

Lower Boundary

The physics of the lower boundary has been developed and extended in various versions of the model, however the focus of the work presented in this thesis is concerned with the dynamics of the upper atmosphere and the details of the lower boundary may be found in Harris (2001). The lower boundary in its simplest form has a fixed temperature of 212.8K and a fixed altitude of 15.525km.

The Poles

Parameters at the poles are set equal to the average values of the adjacent cells. That is, for a given pressure level, all the cells at ±90° are set equal to the average of all the cells at ±88°.

4.2 Solar Energetics

The Sun contributes energy to the terrestrial atmosphere directly via Solar radiation and, more indirectly, via particle precipitation and Joule heating from the Solar wind. In the unmodified version of the model, the relative energy contribution from each of these processes is determined from statistical models that use the F10.7 and K_p indices as proxies.

4.2.1 Solar Radiation

The absorption of radiation in the model is determined using the Beer-Lambert law (2.73), where the optical depth is calculated from (2.74). This is a function of $\sec\chi$, where χ is the angle of incident radiation to the zenith. A result of the Earth's curvature is that radiation originating from a large zenith angle will follow a rather complicated path through the atmosphere and so $\sec(\chi)$ is instead replaced by the Chapman function, $Ch(\chi)$. The following expressions taken from Swider and Gardner (1967) are used to estimate the Chapman function for each species

$$Ch(x,\chi,\leq\frac{\pi}{2})=\left(\frac{\pi x}{2}\right)^{1/2}\left(1-\operatorname{erf}\left(x^{1/2}\cos\frac{\chi}{2}\right)\right)\exp\left(x\cos^2(\chi/2)\right) \tag{4.55}$$

$$Ch(x,\chi,\geq\frac{\pi}{2})=\left(\frac{\pi x}{2}\sin\chi\right)^{1/2}\left(1+\operatorname{erf}\left(-\cot\chi\left(\frac{x\sin\chi}{2}\right)^{1/2}\right)\right)\left(1+\frac{3}{8x\sin\chi}\right) \tag{4.56}$$

where $x=(R+z)/H$ and R is the Earth's radius, z altitude and H the scale height of the absorbing species.

F10.7 Index

Due to variations in solar irradiance at Earth, the value of $I_\infty(\lambda)$ is determined within the model using the F10.7 index as a proxy. This is a daily index that represents the solar radio flux per unit wavelength at a wavelength of 10.7cm.

Solar Heating

The atmospheric heating rate from radiation of wavelength, λ, is given by

$$Q(\lambda)=n_j\frac{hc}{\lambda}\epsilon(\lambda)\sigma_j(\lambda)I_\infty(\lambda)e^{-\tau(\lambda)} \tag{4.57}$$

hc/λ is the energy of a photon of wavelength λ, where h is Planck's constant and c the speed of light. Equation (4.57) is integrated over all wavelengths for all constituent species to calculate the total heating, Q_{SOL}, in the energy equation (4.44).

Solar heating in the model is due to UV between 105 – 180nm and EUV between $1.8-105$nm. The absorption and ionisation cross sections of O, O_2 and N_2 are taken from the tabulated values in Torr et al. (1979). Cross sections for EUV below 5nm are taken from Banks and Kockarts (1973). Solar heating from UV is primarily due to dissociation of O_2 in the range 130 – 240nm, while EUV heating is mainly from ionisation between 3 – 103nm.

4.2.2 Radiative Cooling

The model contains three sources of radiative heat loss within the thermosphere. Firstly, cooling from atomic oxygen, secondly, from NO and thirdly, from CO_2.

O Cooling

Above 110km, radiative cooling is dominated by 63μm emission by atomic oxygen. The model uses the cooling rate from Bates (1951)

$$Q_{IR(O)} = \frac{1.67 \cdot 10^{-25} e^{(-228.0/T_n)}}{1 + 0.6e^{(-228.0/T_n)} + 0.2e^{(-325.0/T_n)}} \cdot [O] \tag{4.58}$$

where T_n is neutral temperature and $[O]$ is atomic oxygen number density. This expression is modified further to avoid overestimation of cooling in the lower thermosphere by the following damping factor

$$Q'_{IR(O)} = Q_{IR(O)} \left(1 - \exp\left(\frac{80 - z}{20}\right)\right)^4 \tag{4.59}$$

where z is altitude in km.

NO Cooling

Between 150-200km, NO 5.3μm emission cooling is accounted for using (Kockarts, 1980)

$$Q_{IR(NO)} = h\nu n\omega A_{10} \left(\frac{g_1}{g_0}\right) e^{-\frac{h\nu}{k_B T}} \tag{4.60}$$

where n is the number of molecules in the vibrational level $\nu = 0$, $h\nu$ is the energy of the photon emitted in the transition from the vibrational level 1 to 0, A_{10} is the probability, and g_0 and g_1 are the statistical weights for the two levels. T is temperature, k_B is the Boltzmann constant and ω is a factor that accounts for the non-Boltzmann distribution of energy levels resulting from the lack of local thermodynamic equilibrium

$$\omega = \frac{k_{10} n(j)}{k_{10} n(j) + A_{10}} \tag{4.61}$$

where k_{10} is the deactivation rate coefficient of NO from collisions with species j and n is the concentration of j. k_{10} is much greater for monatomic oxygen than N_2 and O_2, so only this value is used, where $k_{10}(O) = 6.5 \times 10^{-11}\text{cm}^3\text{s}^{-1}$ at 300K (Fernando and Smith, 1979).

CO_2 Cooling

In the lower thermosphere and upper mesosphere, cooling is mainly due to CO_2 15μm emission. The model uses the parameterisation of the cooling rates from CO_2 by Fomichev et al. (1998).

4.2.3 Electric Field Model

The modifications to CMAT2 detailed in chapter 6 replace the statistical electric field model with real radar observations from SuperDARN. However, the unmodified version of CMAT2 determines energy input from Joule heating and particle precipitation from empirical models using the K_p index as a proxy.

Foster's Electric Field Model

The unmodified version of CMAT2 uses the statistical high latitude electric field model developed by Foster et al. (1986). The intensity and spatial extent of auroral particle precipitation is quantified into 10 discrete levels, each with an associated precipitation index. This index is based on observations made by the TIROS and NOAA satellites (Fuller-Rowell and Evans, 1987). Foster's model uses averaged data selected from five years worth of ionospheric velocity observations by the Millstone Hill radar and bins them according to a TIROS index. A statistical best-fit convection pattern is then derived for each activity level. Assuming the convection is a result of $\mathbf{E}\times\mathbf{B}$ drift, a map of high latitude

K_p	K_p **value**	**TIROS Index**	**Foster field**	$\Delta\Phi$ **(kV)**	E_{GW} **(GW)**
$1-$	0.66	1	1	25.78	4.6
$1\circ$	1.00	2	1	25.78	4.6
$1+$	1.33	3	1	25.78	4.6
$2-$	1.66	4	2	31.83	7.4
$2\circ$	2.00	5	3	34.78	7.4
$2+$	2.33	5	3	34.78	11.7
$3-$	2.66	6	4	46.43	18.7
$3\circ$	3.00	6	4	46.43	18.7
$3+$	3.33	7	5	54.00	29.3
$4-$	3.66	7	5	54.00	45.4
$4\circ$	4.00	8	6	57.05	45.4
$4+$	4.33	8	6	57.05	45.4
$5-$	4.66	9	7	79.87	69.5
$5\circ$	5.00	10	7	79.87	69.5
$5+$	5.33	10	7	79.87	69.5
$6-$	5.66	10	7*	>79.87	>96.0

Table 4.1: K_p levels and their associated Foster fields, cross-cap potentials and particle energy input used by CMAT2. For the highest level, $6-$, the largest Foster field is amplified according to (4.63).

potential can then be associated with each of these convection patterns. Foster's model also includes the average value of the K_p index for each activity level. The model uses one of seven different electric fields, depending on the value of K_p.

The unmodified version of CMAT2 uses the K_p index as a proxy for Joule heating and particle precipitation energies. The model accepts different K_p values, each with an associated Foster high-latitude potential pattern and particle precipitation energy, which are listed in table (4.1). The level $6-$ has a default particle precipitation energy, E_{GW} of 96GW. The user can specify their own energy above this value, which will cause the model to scale up the potential, and therefore the electric field, according to

$$\Delta\Phi(E_{GW}) = 79.87 \cdot \frac{E_{GW}}{96} kV \tag{4.62}$$

For K_p values exceeding $6-$, and if a value is not specified by the user, E_{GW}, is calculated from the equation

$$E_{GW} = ak^2 + bk + c \tag{4.63}$$

where k is the K_p value given in table 4.1 and $a = 3.606$, $b = -3.534$ and $c = 2.487$ (values are taken from CMAT2 code).

To determine the electric field CMAT2 reads in an array of potential values covering 44° to 90° magnetic latitude for the appropriate Foster level. This array is converted to geographic coordinates over both the North and South poles. The model then calculates the x and y components of the electric field from (4.64) using a forward difference scheme. This means the gradient at a given grid point is simply the difference to the next grid point, divided by the spacing between grid points. The potential gradient is assumed to be uniform in the vertical direction

$$\mathbf{E} = -\boldsymbol{\nabla}\Phi \tag{4.64}$$

4.2.4 Particle Precipitation Model

For each K_p level, the associated TIROS level determines the particle precipitation pattern. The total ionisation due to precipitating particles is then determined from the particular pattern. The ionisation rates from the major constituents are determined using the following branching ratios from Jones and Rees (1973)

$$Q_{N_2} = \frac{0.92[N_2]}{\Delta} Q_T \tag{4.65}$$

$$Q_{O_2} = \frac{[O_2]}{\Delta} Q_T \tag{4.66}$$

$$Q_O = \frac{0.56[O]}{\Delta} Q_T + 0.5 Q_{O_2} \tag{4.67}$$

where Q_j is the total ionisation rate for constituent j, Q_T the total ionisation rate due to precipitating particles, $[j]$ the number density of j and Δ is given by

$$\Delta = 0.92[N_2] + 1.5[O_2] + 0.56[O] \tag{4.68}$$

4.3 Chemistry

The chemistry scheme in the model is based upon the schemes of Allen et al. (1984) for oxygen and hydrogen, Solomon et al. (1985) for CO and CO$_2$, and Garcia and Solomon (1983) for odd nitrogen reactions in the mesosphere. The species for which the model solves are given in table 4.2, whilst full tables of the chemical scheme can be found in the theses of Harris (2001) and Dobbin (2005).

Transported	**Partitioned**
Major Constituents	
N_2	
O_2	
O_x	$O = O(^1D) + O(^3P), O_3$
Minor Constituents	
$N(^2D)$	
$N(^4S)$	
NO_x	NO, NO_2
HO_x	H, HO, HO_2
H_2O	
H_2	
CO_2	
CO	
He	
CH_4	

Table 4.2: Constituents solved for in CMAT2.

Atmospheric composition for the three major species within the model is determined using the continuity equation (4.40) in the form (Fuller-Rowell, 1984)

$$\frac{\partial \chi_j}{\partial t} = \frac{1}{\rho} m_j S_j - \mathbf{V}_p \cdot \nabla_p \chi_j - w \frac{\partial}{\partial P} \chi_j - \frac{1}{\rho} \nabla_p \cdot (n_j m_j \mathbf{C}_j) + \frac{1}{\rho} \nabla \cdot (D_{Eddy} n \nabla m \chi_j) \quad (4.69)$$

where $\chi_j = n_j m_j / \rho$ is the mass mixing ratio of species j, S_j represents chemical sources and sinks, m_j molecular mass, n_j number density, $\mathbf{C}_j$ molecular diffusion velocity and D_{Eddy} the eddy diffusion coefficient. m represents the mean molecular mass of the three major constituents

$$m = \frac{n_1 m_1 + n_2 m_2 + n_3 m_3}{n} \quad (4.70)$$

For each constituent, the diffusion velocity, $\mathbf{C}_j$, is found by solving the diffusion equation for a multiple species, non-uniform gas

$$\frac{1}{n} \sum_{i \neq j} \left(\frac{\chi_i}{m_i D_{ij}} n_j m_j \mathbf{C_j} - \frac{\chi_i}{m_j D_{ij}} n_i m_i \mathbf{C}_i \right) = \nabla_p \chi_i + \frac{\chi_i}{m} \nabla_p m + \left(1 - \frac{m_i}{m} \right) \frac{\chi_i}{P} \nabla P \quad (4.71)$$

Similarly to major species, the constituent composition for minor species is solved using a form of the continuity equation but neglecting the transport effects of minor constituents on to the major constituents,

$$\frac{\partial \chi_j}{\partial t} = \frac{1}{\rho} m_j S_j - \mathbf{V}_p \cdot \nabla_p \chi_j - w \frac{\partial}{\partial P} \chi_j + \frac{1}{\rho} \nabla_p ((D_{Eddy} + D_{Molec}) n \nabla_p m \chi_j) \quad (4.72)$$

4.4 Model Parameters

4.4.1 Coefficients of Conductivity

The coefficient of molecular heat conduction, K_m, in the energy equation (4.44) is obtained from the empirical expression (Rees, 1989)

$$K_m = K_0 T^s \quad (4.73)$$

The values of K_0 and s are specific to each of the constituents. For use in equation (4.73) a weighted mean is taken from the three major constituents, O, O_2, N_2, based upon number density. The coefficient of turbulent conductivity is calculated from

$$K_t = c_p \rho D \tag{4.74}$$

where D is the eddy or turbulent diffusion coefficient, ρ mass density and c_p specific heat capacity at constant pressure.

4.4.2 Coefficients of Viscosity

The expression used to calculate the coefficient of molecular viscosity, μ_m, in the momentum equation (4.41) is taken from Dalgarno and Smith (1962)

$$\mu_m = 4.5 \cdot 10^{-5} \left(\frac{T}{1000} \right)^{0.71} \tag{4.75}$$

The coefficient of turbulent heat conductivity is then used to determine the coefficient of turbulent viscosity, μ_t, using

$$\mu_t = \frac{2K_t}{c_p} \tag{4.76}$$

4.4.3 Coefficients of Diffusion

The eddy diffusion coefficients from Reber (1974) and the molecular diffusion coefficients from Colegrove et al. (1966) are given in tables 4.3 and 4.4, respectively.

Altitude [km]	D_{Eddy} **[m^{2}s^{-1}]**
$80 \leq z \leq 105$	100exp[3.0·(z-105)·10^{-5}]+ 50exp[3·(z-105)^2 ·10^{-8}]
$105 \leq z \leq 150$	150exp[-5·(z-105)^2 ·10^{-8}]
$150 < z$	0

Table 4.3: Eddy diffusion coefficients

i-j	D_{ij} $[\mathbf{m^2 s^{-1}}]$
$O - O_2$	$[T/273.0]^{1.5} \cdot [1.01325 \cdot 10^5/p] \cdot 2.60 \cdot 10^{-5}$
$O - N_2$	$[T/273.0]^{1.5} \cdot [1.01325 \cdot 10^5/p] \cdot 2.60 \cdot 10^{-5}$
$O_2 - N_2$	$[T/273.0]^{1.5} \cdot [1.01325 \cdot 10^5/p] \cdot 2.81 \cdot 10^{-5}$

Table 4.4: Mutual molecular diffusion coefficients, where p is pressure in Pascals.

4.5 The Global Ionosphere-Plasmasphere Model

The ionospheric component of the model originates from a coupling between the old UCL thermosphere model (Fuller-Rowell and Rees, 1980) and the Sheffield University model of mid- and high-latitude ionospheric convection (Quegan et al., 1982). This was developed further by Millward et al. (1996a), who created a model based on those of Bailey (1983) and Rippeth (1992). The first of these solves for the densities, temperatures and velocities of H^+ and O^+ ions that are constrained to single flux tubes under the influence of a simple dipole magnetic field. The second model includes a more realistic eccentric dipole approximation, that is, a dipole field tilted relative to the Earth's rotational axis and offset from the Earth's centre.

The motivation for developing a low-latitude model of the ionosphere is that the nature of the geomagnetic field becomes more complex than the high-latitude model was designed for. The paths of the field lines through the atmosphere become more curved at lower latitudes and they are closed, resulting in inter-hemispheric coupling. The global model solves for multiple flux tubes concurrently, each defined by constant magnetic longitude and by an L value, which is the equatorial crossing height in units of Earth radii.

4.5.1 Coordinate System

Coordinate Transformations

The eccentric dipole is chosen over the true IGRF because, although it omits some details of the true structure of the geomagnetic field, it permits the use of a simple coordinate system. In order to relate the geographic coordinates of the model grid to the eccentric dipole coordinate system, a series of transformations are required. Firstly a rotational transformation is applied to the geographic spherical coordinates, (r_g, θ_g, ϕ_g), to give centred dipole coordinates, (r_{cd}, θ', ϕ'), where the dipole has the same orientation as the required eccentric dipole.

$$r_{cd} = r_g \tag{4.77}$$

$$\theta' = cos^{-1}(\cos\theta_n \cos\theta_g + \sin\theta_n \sin\theta_g \cos(\phi_g - \phi_n)) \tag{4.78}$$

$$\phi' = \sin^{-1}\left(\frac{\sin\theta_g \sin(\phi_g - \phi_n)}{\sin\theta'}\right) \tag{4.79}$$

where θ_n and ϕ_n denote the latitude and longitude of the centred dipole axis in geographic coordinates. The centre of this coordinate system is then shifted via a translational transformation given by Fraser-Smith (1987) to obtain an offset dipole in spherical coordinates, (r_m, θ_m, ϕ_m),

$$r_m = ((r_g \sin\theta' \cos\phi' - x_0')^2 + (r_g \sin\theta' \sin\phi' - y_0')^2 + (r_g \cos\theta' - z_0')^2)^{1/2} \tag{4.80}$$

$$\theta_m = \tan^{-1}\left(\frac{((r_g \sin\theta' \cos\phi' - x_0')^2 + (r_g \sin\theta' \sin\phi' - y_0')^2)^{1/2}}{r_g \cos\theta' - z_0'}\right) \tag{4.81}$$

$$\phi_m = \tan^{-1}\left(\frac{r_g \sin\theta' \sin\phi' - y_0'}{r_g \sin\theta' \cos\phi' - x_0'}\right) \tag{4.82}$$

where (x_0', y_0', z_0') is the location of the eccentric dipole centre in centred dipole Cartesian coordinates. The eccentric dipole coordinates, (p, q, ϕ), are then obtained from the following transformations

$$p = \frac{r_m}{r_0}\frac{1}{\sin^2\theta_m} \tag{4.83}$$

$$q = \left(\frac{r_0}{r_m}\right)^2 \cos\theta_m \tag{4.84}$$

$$\phi = \phi_m \tag{4.85}$$

where r_0 is the Earth's radius. p is identical to the L value and has a constant value along a given flux tube. q is orthogonal to p and ϕ and, for a given point on a flux tube, will remain constant under the influence of an $\mathbf{E} \times \mathbf{B}$ drift. The magnetic field used by CMAT2 is the 10^{th} generation IGRF, which is described in section 3.3.3.

4.5.2 Field Line Grid

Each flux tube is defined as a one dimensional grid of points with two base points at a fixed height of 130km. The influence of $\mathbf{E} \times \mathbf{B}$ drift can cause a given flux tube to acquire a velocity vertically, $U_\perp$, or in the East-West, U_ϕ, direction, which may require more grid points to be added due to the nature of the eccentric dipole coordinate system. Each flux tube parameter is therefore defined on an array of size 401, which includes *spare* array elements below 130km that can be used if required.

The global model consists of a large number of flux tubes in order to provide necessary resolution in the geographic coordinate system. There are 300 flux tubes, consisting of 15 *bunches* every 18° of longitude. Within each bunch, the flux tubes are separated in L value between 3.0 and 15.0 corresponding to approximately 200km and 600km.

4.5.3 E×B Drift

Each flux tube moves under the influence of $\mathbf{E} \times \mathbf{B}$ drift with a velocity of

$$\mathbf{U}_{em} = \frac{\mathbf{E} \times \mathbf{B}}{B^2} \tag{4.86}$$

The field-aligned ion velocity, $\mathbf{U}_{||}$, is assumed to be decoupled from $\mathbf{U}_{em}$ so that the ion velocity can be written as $\mathbf{U} = \mathbf{U}_{em} + \mathbf{U}_{||}$. The value of $\mathbf{U}_{em}$ can be determined at every point along the flux tube by calculating it's components, $U_\perp$ and U_ϕ, in magnetic coordinates.

It is found that, as the state of the model evolves, the flux tubes acquire a net drift and move away from their initial positions. This has an adverse effect on spatial resolution and is avoided by forcing each flux tube to return to it's initial position after a given number of time steps by making changes to the value of $U_\perp$ and setting U_ϕ to zero.

Chapter 5

Heliospheric Tomography using HI

This chapter presents results of the tomographic reconstruction of electron density in the solar wind, which is based on many of the techniques detailed in section 2.3. These techniques, usually applied to coronagraph observations near the solar surface, are instead applied to two Earth-impacting CMEs, observed by the wide-angle Heliospheric Imagers. For each of these events a two-dimensional map of electron density in the ecliptic is produced from a series of successive HI-1 images, during the time that each CME remains in the field of view. This is typically a period of approximately 24 hours, allowing an estimate of how the density and position evolve over time as the CME propagates through the heliosphere. This permits an estimate of the CME speed and radial density profile. These are then extrapolated to $1AU$ in order to estimate the densities and velocities that would be observed at Earth, which are compared to measurements from the Wind spacecraft.

This chapter begins with a mathematical discussion of how the Tomographic inversion of HI data is used to determine electron density in the ecliptic. This is followed by a description of how the STEREO images are processed in order to separate the K-corona from the F-corona and background starlight, and the necessary corrections that must be made to the data in order to perform the tomographic inversion. Finally, the selection criteria for CMEs from the available STEREO data is explained. The results of the tomographic reconstruction of two CMEs are presented with their estimated density profiles and velocities. Lastly predictions of their observed values at Earth are made, which are compared to the values measured in-situ at L1, to assess the accuracy of the method.

5.1 The Inverse Problem

The unique viewpoints provided by the STEREO mission allow remote sensing estimates of density structures in the solar wind between the Sun and the Earth. By combining white light observations from the Heliospheric Imagers on board each of the STEREO spacecraft it is possible to estimate the three-dimensional electron density distribution via tomographic reconstruction. Two viewpoints are insufficient to infer detailed CME morphology using tomography, however, when used to estimate density distributions it is possible to arrive at a physically consistent result. CMEs are identified using the heliospheric imagers and densities are then estimated by fitting integrated line-of-sight intensities, measured by each spacecraft, to a 3D heliocentric grid based on Thomson scattering of photospheric light by solar wind electrons. The solution to finding heliospheric densities from HI observations lies in solving the inverse problem

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{\epsilon} \tag{5.1}$$

where $\mathbf{y}$ is the data vector, $\mathbf{x}$ the unknown density vector, $\boldsymbol{\epsilon}$ the noise associated with the data and $\mathbf{H}$ is a measurement operator containing the physics associated with the scattering process. If (5.1) is applied to the density distribution close to the ecliptic plane using HI-1, then $\mathbf{y}$ contains $M = 2048$ elements corresponding to 2 rows of 1024 pixels from each spacecraft. A three-dimensional grid of N elements is chosen, over which $\mathbf{x}$ is defined. $\mathbf{H}$ is then an $M \times N$ matrix, where each element, H_{mn}, is the contribution of a single electron in grid cell n to the intensity in pixel m,

$$H_{mn} = \frac{\pi \sigma_e R_\odot^2}{2 r_0 \sin \epsilon_m} \left(1 - \frac{u}{3}\right) (2 - \sin^2(\phi_n + \epsilon_m)) d\phi_n \tag{5.2}$$

Otherwise, if line-of-sight m does not cross cell n, $H_{mn} = 0$. The dot product $\mathbf{Hx}$ gives an array where each element is a finite difference approximation to the integral in (2.69),

$$(\mathbf{Hx})_m = \frac{\pi \sigma_e R_\odot^2}{2 r_0 \sin \epsilon_m} \left(1 - \frac{u}{3}\right) \sum_{n}^{n_{max}} (2 - \sin^2(\phi_n + \epsilon_m)) N_{en} d\phi_n \tag{5.3}$$

5.1.1 Model Grid

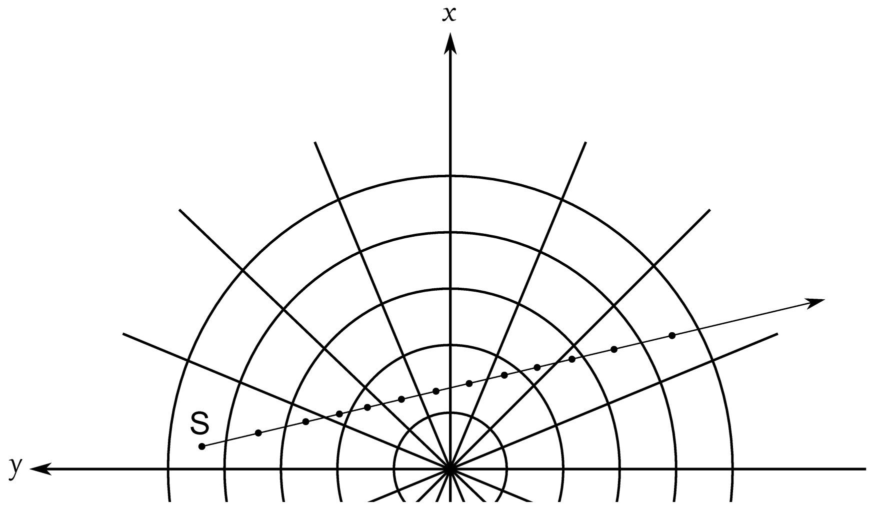


Figure 5.1: The grid over which the elements of matrix **H** are computed. S represents either STEREO spacecraft and the dots indicate where the scattering is calculated in each cell. The x and y axes correspond to the HEE coordinate system, that is, Earth lies on the x-axis and the x-y plane is the ecliptic.

The coordinate system used for computation of the densities is similar to that of Tappin and Howard (2009b); a three-dimensional, heliocentric, spherical grid with resolution $0.025\mathrm{AU}\times 3^\circ \times 3^\circ$. The radial limit of the grid is 2AU, resulting in $> 10^6$ grid cells. However, cells that are not common to both HI-1 fields of view are ignored and we are only working in the ecliptic plane. So, in practice, there will be only several thousand non-zero elements to $\mathbf{x}$.

An algorithm, similar to that used by Tappin and Howard (2009b) for calculating (2.65), is applied to the problem of determining the elements of **H**. The value of (5.2) is determined for every cell crossed by each line-of-sight, up to a given distance, $z_{max} = 2(1+\sin(\epsilon/2))$AU. The mathematics of this process is described in detail in appendix A, however the basic algorithm takes the following form

1. Starting at the observer, the distance to the next cell boundary is computed in each coordinate.
2. The distance through the cell is the shortest of these.
3. Compute H_{mn} from (5.2) at the mid-point of the cell.

4. Repeat steps 1-3, starting from the next cell.

5. Terminate when $z \geq z_{max}$.

Once the algorithm has been completed for all lines-of-sight, the matrix **H** is found to contain a large number of zeroes. If the n^{th} grid cell lies outside the STEREO field of view, the n^{th} column of **H** will be empty as a consequence. For example, as illustrated in (5.4), if no line-of-sight crosses cells 1, 2 and 4, the corresponding columns will contain all zeroes.

$$\begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_M \end{pmatrix} = \begin{pmatrix} h_{0,0} & h_{0,1} & h_{0,2} & h_{0,3} & h_{0,4} & \cdots & h_{0,N} \\ h_{1,0} & h_{1,1} & h_{1,2} & h_{1,3} & h_{1,4} & \cdots & h_{1,N} \\ h_{2,0} & h_{2,1} & h_{2,2} & h_{2,3} & h_{2,4} & \cdots & h_{2,N} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ h_{M,0} & h_{M,1} & h_{M,2} & h_{M,3} & h_{M,4} & \cdots & h_{M,N} \end{pmatrix} \cdot \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ \vdots \\ x_N \end{pmatrix} \tag{5.4}$$

The matrix is then reduced by omitting these redundant columns, and the corresponding elements of **x**, leaving only non-zero columns,

$$\begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_M \end{pmatrix} = \begin{pmatrix} h_{0,0} & h_{0,3} & \cdots & h_{0,N} \\ h_{1,0} & h_{1,3} & \cdots & h_{1,N} \\ h_{2,0} & h_{2,3} & \cdots & h_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ h_{M,0} & h_{M,3} & \cdots & h_{M,N} \end{pmatrix} \cdot \begin{pmatrix} x_0 \\ x_3 \\ \vdots \\ x_N \end{pmatrix} \tag{5.5}$$

The dimensions of array **x** are reduced from $N \sim 10^6$ to $\sim 10^3$, containing only grid cells that are common to the fields of view of *both* STEREO-A and STEREO-B. Compared to the $M = 2048$ elements of array **y**, this makes the problem (5.1) more easily solvable. However the large number of zeroes means that there is unlikely to be a unique solution and the problem cannot be dealt with analytically.

5.1.2 Solution to the Inverse Problem

Gradient Term and Regularisation Parameter

In order to reduce the contribution of noise to the reconstructed density, Frazin (2000) introduces a condition that ensures the reconstructed density is smooth. This is achieved by rewriting the inverse problem in terms of a new matrix, **A**, that satisfies the equation

$$\mathbf{b} = \mathbf{A}\mathbf{x} \tag{5.6}$$

where **b** and **A** are given by

$$\mathbf{b} = \begin{pmatrix} \mathbf{y} \\ \underline{\mathbf{0}} \\ \underline{\mathbf{0}} \end{pmatrix} \tag{5.7}$$

$$\mathbf{A} = \begin{pmatrix} \mathbf{H} \\ \lambda \mathbf{D}_r^2 \\ \lambda \mathbf{D}_\phi \end{pmatrix} \tag{5.8}$$

The array **b** is created by extending the array **y**, with a number of additional elements equal to zero. The matrix, **H**, ensures the estimate, $\hat{\mathbf{x}}$, is consistent with the data and the matrices, $\mathbf{D}_r^2$ and $\mathbf{D}_\phi$, suppress large gradients in the resulting estimate. $\mathbf{D}_r^2$ and $\mathbf{D}_\phi$ are a discrete approximation of the respective gradient operators, $\partial^2/\partial r^2$ and $\partial/\partial\phi$, and the regularisation parameter, λ, can be tuned to control their relative contribution to the solution.

Weighting Factor

Due to the large range in observed intensities (approximately three orders of magnitude in HI-1), (5.1) is multiplied by a weighting factor, w, to increase the contribution from lines-of-sight that lie at large distances from the Sun. Kramar et al. (2009) suggest values of

$$w_m = \frac{1}{y_m - \min(\mathbf{y})/10} \tag{5.9}$$

Both element $\mathbf{y}_m$ and row $\mathbf{H}_m$ in (5.1) are multiplied by w_m, which still delivers the solution, $\mathbf{x}$, corresponding to the original system $\mathbf{Hx} = \mathbf{y}$. The answer to the problem of reconstructing the density array results from minimising the following function

$$\Phi(\hat{\mathbf{x}}) = \|\mathbf{b} - \mathbf{Ax}\|^2 \tag{5.10}$$

That is,

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\mathrm{argmin}}[\Phi(\mathbf{x})|x_i \geq 0] \tag{5.11}$$

Rather than attempting to solve this analytically, the array $\hat{\mathbf{x}}$ is estimated via an iterative process.

Conjugate Gradient Method

The conjugate gradient (CG) method is an iterative algorithm for numerically solving large sets of simultaneous equations[1]. For equations such as this, the number of calculations needed to reach an analytical solution is proportional to N^3. The CG method is a numerical approximation, which can greatly reduce this time. Because the method is iterative it may be stopped after a finite number of steps, once a sufficient level of accuracy has been reached. Several approaches to solving linear problems using iterative methods, including the CG algorithm and its variants, are covered in Barrett et al. (1994). The principle common to all versions of the CG algorithm is that the solution, $\hat{\mathbf{x}}$, may be expanded in terms of a set of basis vectors, $\mathbf{p}_i$,

$$\hat{\mathbf{x}} = \sum_{i=1}^{N} \alpha_i \mathbf{p}_i \tag{5.12}$$

---

[1] The conjugate gradient method in fact refers to a group of methods, each of which may be applicable depending on the nature of the matrix, $\mathbf{A}$, in the inverse equation.

where N is the dimension of $\hat{\mathbf{x}}$. The array $\mathbf{b}$ is then expressed in terms of $\mathbf{p}_i$ by

$$\mathbf{b} = \mathbf{A}\hat{\mathbf{x}} = \sum_{i=1}^{N} \alpha_i \mathbf{A}\mathbf{p}_i \tag{5.13}$$

Provided that $\mathbf{p}$ is a sequence of mutually conjugate[2] vectors, the coefficients may be found by the following relation

$$\mathbf{p}_k^{\mathbf{T}}\mathbf{b} = \mathbf{A}\hat{\mathbf{x}} = \sum_{i=1}^{N} \alpha_i \mathbf{p}_k^{\mathbf{T}} \mathbf{A}\mathbf{p}_i = \alpha_k \mathbf{p}_k^{\mathbf{T}} \mathbf{A}\mathbf{p}_k \tag{5.14}$$

$$\alpha_k = \frac{\mathbf{p}_k^{\mathbf{T}}\mathbf{b}}{\mathbf{p}_k^{\mathbf{T}}\mathbf{A}\mathbf{p}_k} \tag{5.15}$$

The simplest form of the CG algorithm (Hestenes and Stiefel, 1952) is a method of solving (5.12) iteratively for each value of α_k and $\mathbf{p}_k$, although it is only applicable to symmetric, positive-definite matrices. Since neither of these conditions are true of the matrix $\mathbf{A}$ in (5.10), the more recent BiConjugate Gradient Stabilised (BiCGSTAB) algorithm (van der Vorst, 1992) is used instead. This is applicable to non-symmetric systems and has been demonstrated to converge more rapidly than alternative methods. The BiCGSTAB method works by approaching the solution, $\hat{\mathbf{x}}$, in terms of two mutually orthogonal sequences of basis vectors, at the cost of no longer providing an exact minimisation. The pseudocode for the algorithm is shown below (Barrett et al., 1994).

The inclusion of the gradient operators in (5.8) prevent the routine from converging to some extent and, accordingly, the loop is terminated when the residual, $r^{(i)}$, reaches a sufficiently low value. To ensure the resulting densities are positive, any negative values that arise are reset after each iteration using the corresponding value from $x^{(0)}$.

---

[2] Two vectors, $\mathbf{u}$ and $\mathbf{v}$, are said to be conjugate with respect to matrix $\mathbf{A}$ if they satisfy the equation $\mathbf{u}^{\mathbf{T}}\mathbf{A}\mathbf{v} = 0$, and are non-zero.

Algorithm 1 The BiConjugate Gradient Stabilised Method

```
Compute $r^{(0)} = b - Ax^{(0)}$ for initial guess $x^{(0)}$
$\tilde{r} = r^{(0)}$
for $i = 1, 2, \ldots$ do
    $\rho_i = \tilde{r}^T r^{(i-1)}$
    if $i = 1$ then
        $p^{(i)} = r^{(i-1)}$
    else
        $\beta_{i-1} = (\rho_{i-1} - \rho_{i-2})(\alpha_{i-1}/\omega_{i-1})$
        $p^{(i)} = r^{(i-1)} + \beta_{i-1}(p^{(i-1)} - \omega_{i-1}v^{(i-1)})$
    end if
    solve $M\hat{p} = p^{(i)}$
    $v^{(i)} = A\hat{p}$
    $\alpha_i = \rho_{i-1}/\tilde{r}^T v^{(i)}$
    $s = r^{(i-1)} - \alpha_i v^{(i)}$
    solve $M\hat{s} = s$
    $t = A\hat{s}$
    $\omega_i = t^T s / t^T t$
    $x^{(i)} = x^{(i-1)} + \alpha_i \hat{p} + \omega_i \hat{s}$
    $r^{(i)} = s - \omega_i t$
end for
```

5.2 Data Selection

5.2.1 Image Processing

Of the sources of light that contribute to the total intensity in a given HI pixel, only the contribution from the Thomson scattered K-corona may be used to infer electron densities. In addition to this, the image contains an approximately constant contribution from dust-scattered light, known as the *F-corona*, and a contribution from the background star-field. The K-coronal brightness is typically a few percent of the total brightness at elongation angles of a few degrees and this becomes even less when moving away from the Sun (Eyles et al., 2009). For this reason tomographic reconstructions are generally performed using coronagraph images within $10\mathrm{R}_\odot$ of the Sun. Using the methods described in this section it is possible, to some extent, to separate the K-corona and to perform tomography using the STEREO Heliospheric Imagers.

The presence of planets, comets and other bodies during some periods of HI data contribute to the observed intensities via reflected sunlight. There are further contributions that occur due to dust impact signatures (Davis et al., 2012) and internal reflections (Halain et al., 2011). In addition there is an approximately biannual passage of the galactic plane through the field of view, which is considerably brighter than the typical background star-field. These contributions prove more difficult to separate from the K-corona and so render certain periods of HI data unusable for tomography.

Separation of the K-corona

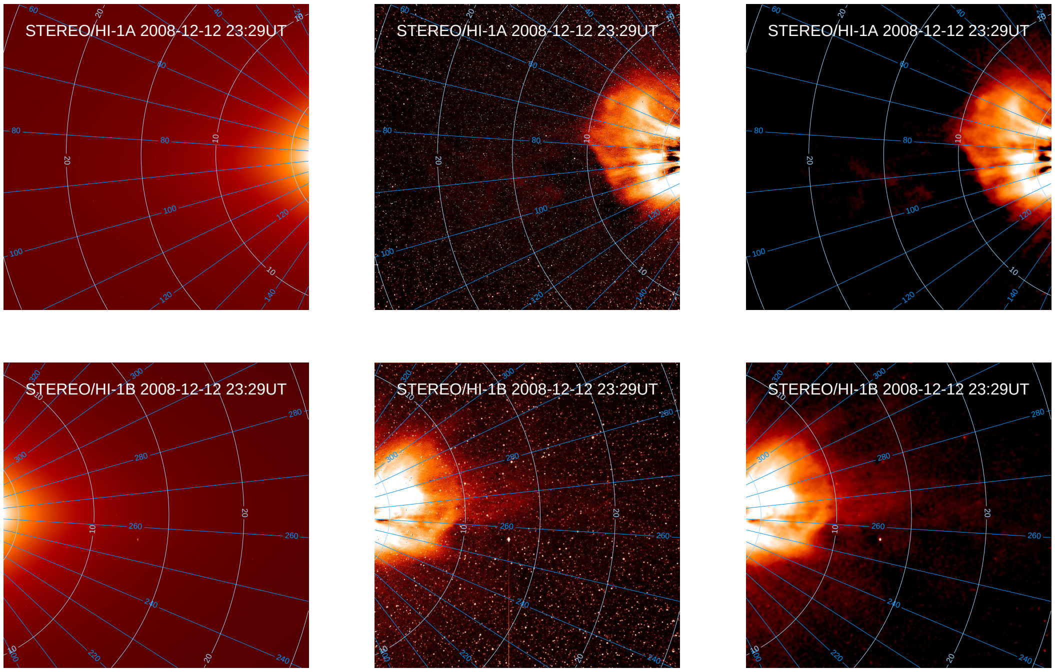


Figure 5.2: STEREO HI-1 images showing the three stages of image processing. The top row contains images from STEREO-A and the bottom from STEREO-B. The plots show the same image at each of three stages of processing, from left to right. The left hand images show L1 data, the middle images show L2 data with a one-day minimum subtraction and the right hand plots have had a median filter applied to remove stars.

The raw data from each HI undergo three stages of image processing to extract the K-corona. Firstly the data are corrected for a large-scale flat-field (Bewsher et al., 2010) and shutterless readout (Eyles et al., 2009). The flat-field correction accounts for uneven responses of pixels across the CCD array by applying a gain-map, derived from analysing the star-field (Bewsher et al., 2010). A pointing calibration is applied to the data, by comparing the locations of stars in the HI image with their positions predicted using a star catalogue (Brown et al., 2009). The finite time taken to read and clear the CCD array results in a known vertical smearing of the image and is accounted for mathematically. All of these processes are standard calibration techniques performed in order to produce the L1 from the L0 data.

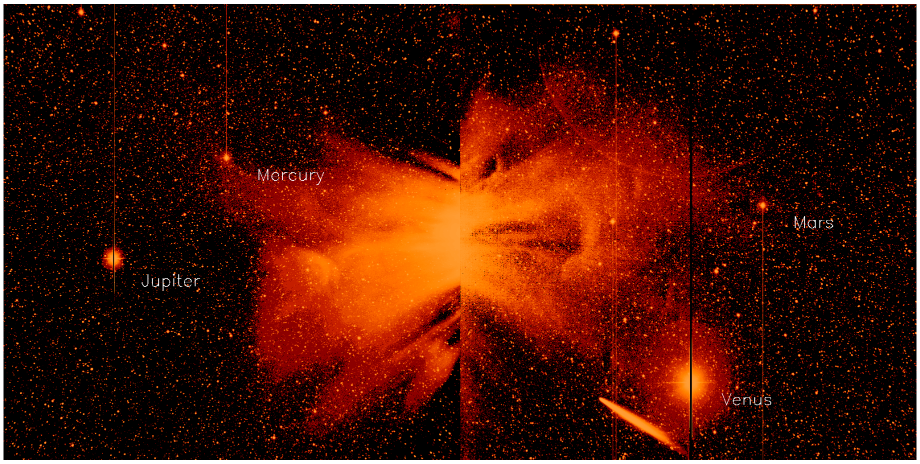
The second stage of processing is to remove the contribution from the F-corona and thirdly, and lastly, the contribution from the background star-field is subtracted. Figure 5.2 shows the results of each of these stages of processing, for a single pair of images. The second and third stages of processing, as performed using the methodologies adopted here, each depend on two parameters. If we assume that the F-corona does not vary on time scales shorter than several days, it may be removed using the following method. A sequence of images is taken over a period of m days. This nominally consists of 36 images per day for HI-1. The F-corona may then be approximated, for each individual pixel, as average of the lowest n^{th} percentile of all values in that pixel. These values may then be subtracted from an individual image to reveal an estimate of the K-corona. For example, in figure 5.2 the minimum value from 1 day of data, centred on each image, is removed to produce the middle two images, i.e. $m = 1$ and $n = 0$. This processing is performed at RAL Space, who are responsible for the HI instruments, and is made available as L2 data.

Finally, the star-field must also be removed. This is achieved, here, by replacing each pixel by the robust mean of its neighbourhood; from an array of $p \times p$ pixels centred on each image pixel the median and median absolute deviation are calculated. Any pixel values exceeding q deviations from the median are excluded and the mean value of the remaining pixels is returned. This procedure is used to exclude outliers that are presumed to result from background stars. In our case, $p = 20$ and $q = 2$ are used to produce the final (right hand) images in figure 5.2.

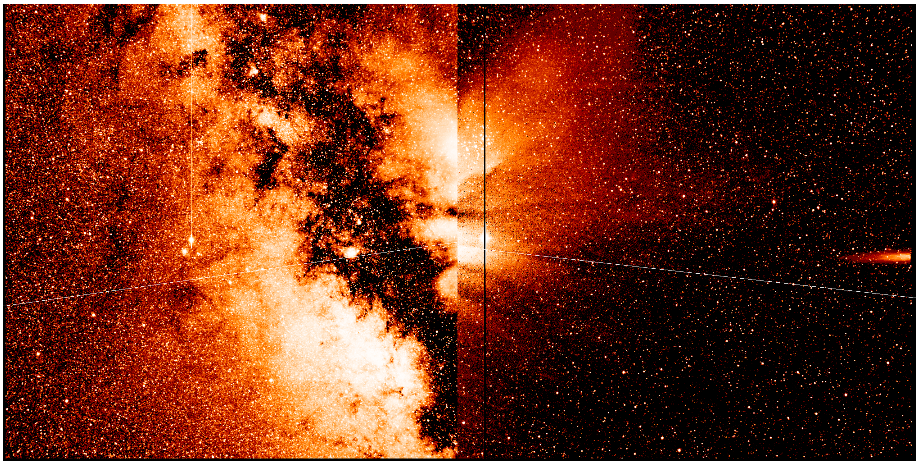
After the launch of STEREO, it was found that the pointing of the HI instruments on STEREO-B made many small movements. This was determined to be due to the impact of dust particles, due to the fact that the cameras on STEROE-B are facing its direction of motion (Davis et al., 2012). The scale of these movements is very small and translates to $\sim \frac{1}{3}$ of a pixel. It is however enough to severely affect the background subtraction and typically causes inconsistencies between the F-corona subtraction on STEREO-A and STEREO-B.

Solar System Bodies and Other Objects

Sunlight reflected by planets and comets within the solar system produces images such as figure 5.3(a). Unlike the optically thin F-corona, the light from objects within the solar system cannot easily be removed using image processing. CCD pixels become saturated by light from planets, comets and the brightest stars, causing the signal to spill over into adjacent pixels, appearing as a bright column in the image (Eyles et al., 2009). Depending on their distance from the observing spacecraft, cometary tails can affect large areas of the image array and are near impossible to remove. These objects generally appear in, or near, the ecliptic plane, making tomography in this region very difficult.



(a) HI-1 data containing several planets in the fields of view of both spacecraft and a cometary tail seen in ST-B. In addition an internal reflection can be seen in the top of the ST-B image.



(b) HI-1 data, where the galactic plane almost covers the field of view of ST-A. An internal reflection is again visible on the right hand side of the ST-B image.

Figure 5.3: Two examples of HI-1 data, with the F-corona removed, that are unsuitable for estimating electron densities. Whilst the bottom right image is only affected by a small internal reflection, the tomographic inversion requires both images in order to arrive at a result. The images are the same format as figure 5.2, with ST-A on the left and ST-B on the right.

Whilst the background star-field can, to some extent, be removed using the image processing technique described above, the density of stars in the galactic plane means this is not possible. Figure 5.3(b) shows an image of the galactic plane in HI-1A, which passes through each HI field of view approximately twice per year. The difficulty in removing this makes tomography impossible while it is present in an image.

Instrument Sensitivity

Accurate reconstruction of electron densities depends on the intensities observed by HI, and is ultimately limited by the brightness sensitivity of the instruments. The values ($3 \times 10^{-15} B_\odot$ and $3 \times 10^{-16} B_\odot$ for HI-1 and 2 respectively, table 3.3) are over plotted in figure 5.4 for comparison with the intensity of the K-corona and star-field. It can be seen that, towards the edge of the field of view in HI-1 and 2, the K-corona becomes comparable to the instrument sensitivity. Additionally, in HI-2, the K-corona is overwhelmed by starlight entirely.

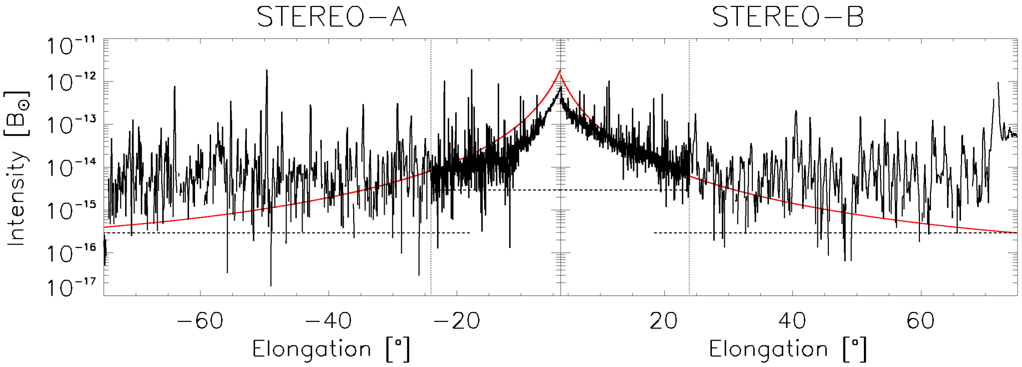


Figure 5.4: The ecliptic intensity seen in both cameras on both spacecraft. The vertical dotted lines indicate the edge of the HI-1 field of view and the horizontal dashed lines represent the brightness sensitivity of each instrument. The red lines represent the radiance that we would expect to observe from a typical electron density profile, which is explained in section 5.3.

From the derivations in chapter 2, the observed intensity in a region of the heliosphere is known to be a decreasing function of the distance of that region from the Sun. This would imply an upper limit for the heliocentric distance at which tomographic reconstruction of CMEs can be performed. For even the densest CMEs, the observed brightness will become less than that of the background star-field as they travel away from the Sun. This means that tomography is only effective at small heliocentric distances and typically not in the HI-2 field of view. The inversion technique will therefore be performed using only the HI-1 data, a region that does not contain the Earth until later in the mission where the spacecraft separation is unsuitable. Estimates of CME velocities and densities must then be extrapolated if we wish to estimate properties of the solar wind in the near-Earth environment.

5.2.2 Spacecraft Configuration

On the order of thousands of CMEs have occurred since the launch of the STEREO mission, many of them passing through the fields of view of its cameras. The events to which the tomographic inversion may be applied is, however, restricted to only a small number due to several constraints. Firstly, the method may only be applied to the region of the heliosphere that is common to the field of view of *both* spacecraft, through which a CME must pass in order to permit tomographic inversion. From a total of 24 CMEs identified by Möstl et al. (2014) 19 were directed toward Earth, many of which must be discarded here due to the orbital configuration of the STEREO spacecraft. This is because when the spacecraft are separated by $\sim 180°$, the HI-1 lines-of-sight become approximately parallel to each other and therefore each spacecraft receives approximately the same projection information, making tomography ineffective. The intersection of the STEREO lines-of-sight also results in some assumptions that must be made in order to apply the method to STEREO data, regarding the contribution to each pixel from light scattered in the regions *not* common to both spacecraft.

The tomographic inversion is only applicable to the region common to both HI fields of view, region 2 in figure 5.5, and so CMEs may only be studied whilst they remain in this region. It is therefore necessary to make a decision about how the contribution to the total brightness from regions 1 and 3 may be removed. This is achieved by assuming a smooth radial density profile, which is discussed in the next section, and estimating what contribution to the total intensity originates in region 2 (figure 5.5). This fraction of the total intensity is then applied to the data as a correction factor.

5.3 Density of the Quiet Corona

5.3.1 Radial Density Profile of the Solar Wind

The tomographic reconstruction of electron densities is the solution to the inverse problem, (5.1),

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{\epsilon} \tag{5.16}$$

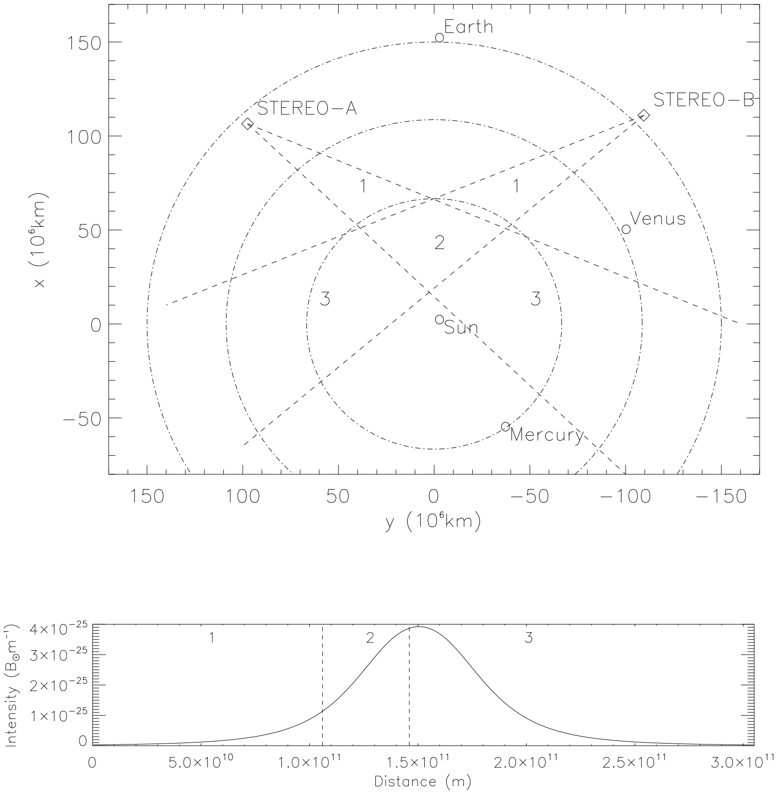


Figure 5.5: The top plot shows the configuration of STEREO-A and -B and the inner planets in the ecliptic plane on 12^{th} December 2008. The fields of view of HI-1 on each spacecraft are indicated by the dashed lines, where the region labelled “2” is common to both instruments. The bottom plot shows the relative contribution to the intensity observed in a central pixel from each of the regions 1, 2 and 3.

where $\mathbf{y}$ is the STEREO data, $\mathbf{x}$ the density and $\mathbf{H}$ an operator containing the physics of Thomson scattering. We can be confident from the derivations in chapter 2 that the operator $\mathbf{H}$ contains the correct physics, so the form of $\mathbf{x}$ then depends upon how we manipulate the image, i.e. on $\mathbf{y}$. For any equation of the form of (5.1), there exist values of $\mathbf{y}$ and $\mathbf{H}$ that have no mathematical solution, where the problem is called *inconsistent*. When applied to solar wind densities, we must satisfy the constraints that the densities are positive and, according to the regularisation terms $\mathbf{D}_\phi$ and $\mathbf{D}_r^2$ in (5.8), they must have a smooth gradient. This means the image manipulation techniques described earlier in this chapter *must* produce a value of $\mathbf{y}$ that is both mathematically and physically consistent or the inverse problem cannot be solved. The background subtraction method depends on four parameters, which may be tested to some extent with prior knowledge of the radial electron density profile in the heliosphere. Several authors have attempted to measure the density profile using both direct and indirect methods, however the results (table 5.1 and figure 5.6) do not generally agree, or they are inconsistent with the region of the heliosphere of interest here.

Baumbach (1937) was the first to attempt this calculation using the quantitative expressions of Minnaert (1930) for coronal line-of-sight brightness. The author assumed a spherically symmetric coronal density, with a dependence on r. The expression results in densities away from the Sun that are typically much lower than more recent measurements. In addition the author does not quote any of the uncertainties that result from the method.

Radio scintillation measurements from Mariner 6 and 7 were used by Muhleman et al. (1977) to estimate the coronal density profile. The light travel time of the spacecraft transmissions are affected by general relativistic effects near the Sun, but there is an additional delay caused by the electron content along the path of the signal. The authors determined a power law for the electron density profile, although their method results in very large uncertainties in their coefficient and index.

Most recently, Gazis et al. (2006) have used in situ measurements from spacecraft at varying latitudes and altitudes throughout the heliosphere. Measurements from Helios 1 & 2, ACE, WIND, Ulysses and Voyager 2 are bin-averaged according to the presence or absence of CMEs, providing a density profile fitted to a power law for both the active and quiet solar wind, respectively. This is however based on measurements distributed throughout the heliosphere, and not necessarily representative of the region of interest here; the ecliptic plane inside 1 AU. Due to the lack of a definitive estimate of the density profile in this region, a profile has been fitted using bin-averaged density measurements made *only* by the Helios spacecraft, which operated in ecliptic orbits within 1 AU.

Author	**Density Profile**
Baumbach (1937)	$n_e(r) = (5.25r^{-2.5} + 7.08) \times 10^{-9}r^{-7} + 6.16 \times 10^{-32}r^{-17}$
Muhleman et al. (1977)	$n_e(r) = (9 \pm 7)r^{-2.08\pm0.23}$
Gazis et al. (2006)	$n_e(r) = 7.96 \pm 0.38r^{-2.08\pm0.03}$

Table 5.1: Measurements of the radial number density profile of the solar wind produced using several techniques. The densities are quoted per cc, with r in AU.

Helios 1 and 2 were launched in to eccentric heliocentric orbits with respective mission durations of 1975-1985 and 1976-1979. Their orbital radii varied between 0.3 and 1.0AU, allowing a measurement of the density profile within this range. Figure 5.6 shows the

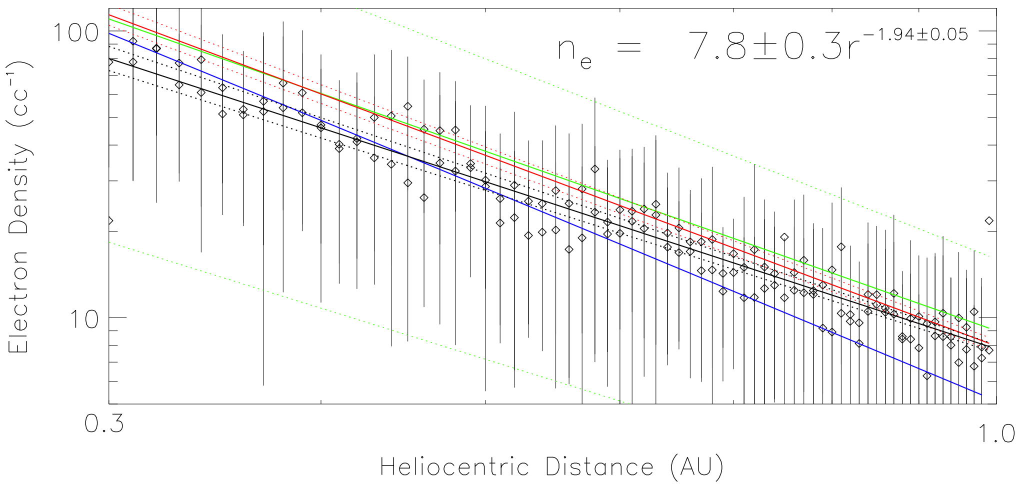


Figure 5.6: The radial density profile between 0.3 and 1AU produced using averaged HELIOS data, shown in black. The values from Baumbach (1937) (blue), Muhleman et al. (1977) (green) and the quiet value from Gazis et al. (2006) (red) are over plotted for comparison. The dotted lines represent the uncertainties in each profile.

profile produced using data for a period of just over three years (∼6 orbits), whilst both spacecraft were simultaneously returning data.

The resulting density profile is a power law given by $n_e(r) = (7.8 \pm 0.3)r^{-1.94\pm0.05}$, which is close to the value quoted by Gazis et al. (2006) for the quiet solar wind, although with a slightly shallower slope resulting in lower densities at small r. Knowledge of this relationship between n_e and r is useful in several aspects of solving the inverse problem. The profile provides a good value for the initial guess, $\mathbf{x}^{(0)}$, in the iterative solution of the inverse problem. Its value is necessary for the use of the regularisation terms, $\mathbf{D}_\phi$ and $\mathbf{D}_r^2$, and for removing the fractional intensity outside the combined field of view, i.e. regions 1 and 3 in figure 5.5. Finally, the density profile provides a means of assessing the effectiveness of the background subtraction of the F-corona and star-field, which is the subject of the following section.

5.3.2 Tomographic Reconstruction of the Quiet Corona

STEREO Observations

The methods for separating the K-corona, F-corona and background star-field in HI data each rely on two control parameters. Subtraction of the F-corona depends on both the period over which images contribute to the subtraction and the percentile that is to be averaged and subtracted. The removal of the star-field depends on both the size of the array used to compute the average and the number of standard deviations outside which data are ignored. The latter two are found to have a much less significant effect on the result and so it is important to assess which of the F-corona subtraction methods produces a realistic K-corona intensity.

We know that on average the density of the K-corona follows the relation $n_e(r) = (7.8 \pm 0.3)r^{-1.94\pm0.05}$, however we cannot say what its value will be at a given instant, i.e. in a single HI image. If, however, we take a suitably large number of consecutive images and average each pixel, we can assume with confidence that the intensity results from a density distribution that follows the power law. This provides a means to test the effectiveness of the different parameters in subtracting the F-corona.

A two-week period from 14 – 27 December 2008 was chosen where solar activity was relatively quiet and the HI fields of view were at approximately 90°. For three combinations of parameters the background is removed by subtracting the minimum over one day, the average of the lowest 25^{th} percentile over one day and the average of the 25^{th} percentile over 11 days. For all images, each pixel value is averaged over the two-week period. The resulting values and their standard deviations[3] are plotted against the radiance that would be expected from $n_e(r) = 7.8 \pm 0.3r^{-1.94\pm0.05}$ electrons. This is calculated by solving $\mathbf{y} = \mathbf{Hx}$ for $\mathbf{y}$, with a known density array (figure 5.7), using the relationship above.

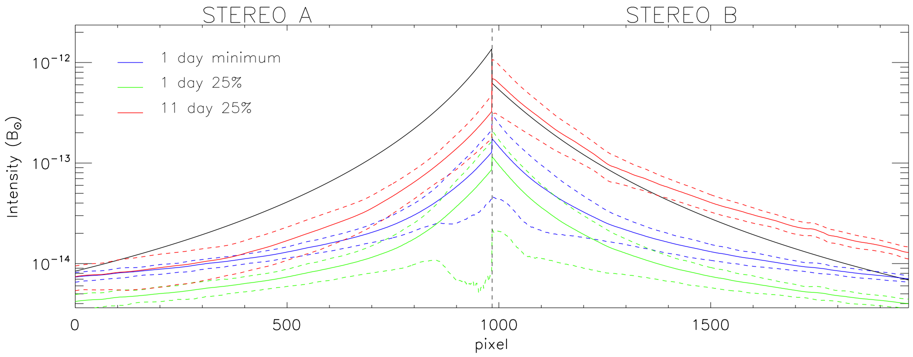


Figure 5.7: The mean ecliptic radiance observed by STEREO HI-1 over a two-week (14-27^{th} December 2008) period using three sets of L2 data. The solid lines show mean values and the dashed lines show standard deviation. The black line shows the theoretical value from a density of $7.8r^{-1.94}$ electrons cc^{-1}, determined using (5.16).

The resulting plot suggests that in fact none of the three combinations of parameters are good at producing the intensities that correspond to realistic electron distributions.

---

[3] These are calculated in the ecliptic plane only, due to the computational costs of processing two weeks of data.

When using either of these three data sets with the inverse problem it is found that the solution rarely converges to a physically consistent result. The conjugate gradient algorithm evolves towards a solution containing negative densities, which must be replaced with positive values after each iteration. This method is effective if a small number of negative densities occur, however if this number becomes large enough, the solution is prevented from converging entirely. Without more thorough background subtraction techniques, it becomes necessary to apply some correction to the data in order to arrive at a solution.

For an individual image in HI, we cannot know the exact difference between the true K-corona intensity and the estimate. Based on figure 5.7 however, we can say *on average* by how much we under- or over-estimate the K-corona in each pixel. For either of these three data sets we can quantify this by taking the ratio of the observed intensity to the expected intensity in each pixel. This may then be applied to an individual image as a scaling factor.

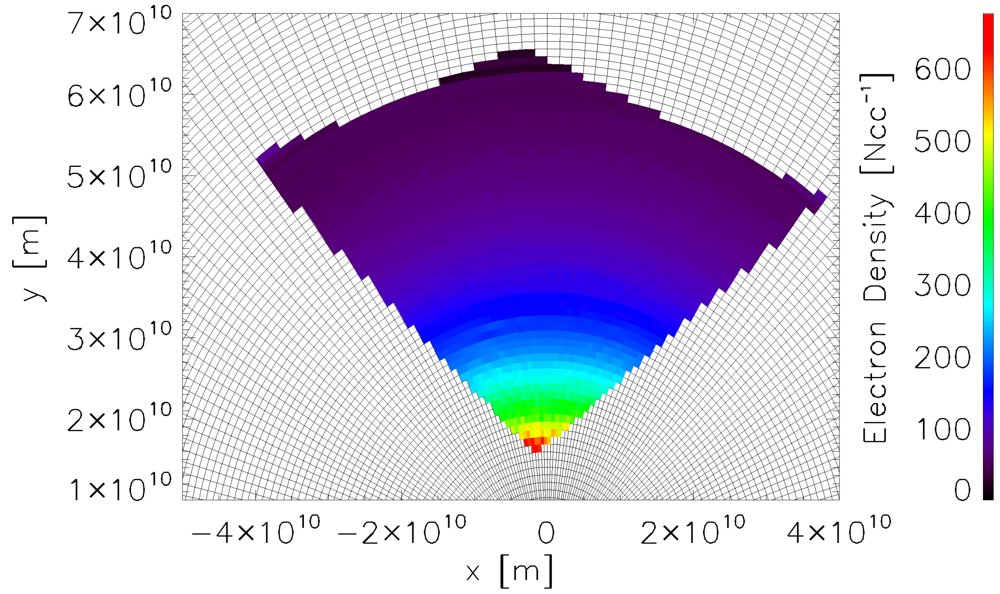
Density Map

This technique may be demonstrated by applying it to a HI data set in order to produce a density estimate. The 28^{th} December 2008 was selected following the two-week period used in the previous discussion. The data[4] are again averaged, this time over the single day, in order to produce a data array that represents a smooth density. The correction factor from figure 5.7 is applied to the data and the resulting density distribution is shown in figure 5.8. This is performed without the use of the radial gradient term in (5.8).

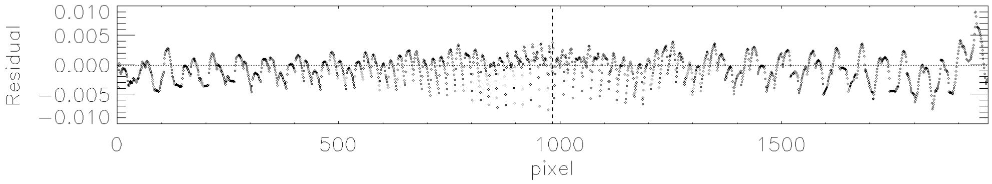
The resulting density along the Sun-Earth line is plotted in figure 5.9, with the profile derived from Helios over plotted for comparison. The resulting profile from the inversion is found to be $n_e(r) = (6.9 \pm 0.6) r^{-2.08 \pm 0.06}$, which is a physically realistic density estimate, based on the Helios observations. Although this result is based on data averaged over one day, rather than on a single image, this demonstrates the effectiveness of the data correction in producing a realistic K-corona. The result also serves to demonstrate the effectiveness of the tomographic inversion, where the residual between the model and data has an RMS value of $\sim 0.2\%$.

---

[4] The one-day-minimum background subtraction method is chosen because it is found to reveal the structure of CMEs most effectively, despite not being the closest fit to the expected intensity.



(a) Electron number density in the region of the ecliptic that is common to the fields of view of both STEREO-A and B.



(b) The relative residual between the data and solution, resulting from the above inversion, which has an RMS value of 2.32×10^{-3}.

Figure 5.8: Results of the tomographic inversion on the data averaged over 36 HI-1 images on 28^{th} December 2008.

5.4 Density Estimates of Coronal Mass Ejections

5.4.1 Event List

According to Möstl et al. (2014), based on data from STEREO, a total of 19 Earth directed CMEs have occurred between the launch of STEREO and the date of publication. Of these events, only two may be studied using tomography due to the limitations involved with the data that are discussed earlier in this chapter. Table 5.2 shows the list of events presented by Möstl et al. (2014), with the usable events highlighted.

Events 6, 7 and 22 are rejected because they occurred whilst the galactic plane was in the HI field of view and events 17-24 whilst planets were present. Events 8-15 are rejected because they occurred during the period when STEREO-A and B were approximately

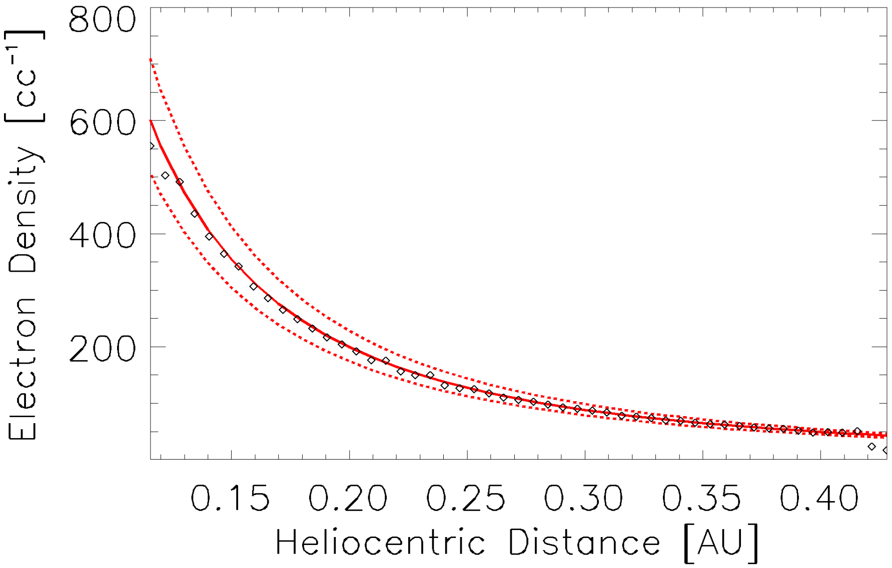


Figure 5.9: The electron density as a function of heliocentric distance along the Sun-Earth line, from the result in figure 5.8. The profile derived from Helios data is over plotted for comparison.

opposite each other, which creates line-of-sight ambiguities in the tomographic inversion.

For the remaining two events, 12 December 2008 and 22 October 2011, the tomographic reconstruction is performed to determine n_e. Each reconstruction is carried out using a series of successive images spanning the time in which the CME remains in the HI-1 field of view.

5.4.2 December 2008 CME

The CME launched on 12 December 2008, which passed over ACE four days later was the first clear Earth-impacting event to be observed during the STEREO mission, eg. Davis et al. (2009). 72 HI-1 images, with daily minimum background subtraction, from both spacecraft were selected covering the 48 hour period beginning at 00:09 on 12 December 2008. The star-removal method, described previously, was applied to the pixels lying in the ecliptic plane, using a grid of 20 × 20 and excluding data that exceeded two standard deviations. Lastly, the correction is applied to the data using the ratio between the expected intensity and the mean observed intensity from the preceding 14 days, to produce $\mathbf{y}$. The elements of the observation matrix, $\mathbf{H}$, are computed for each line-of-sight over a grid with a resolution of 1/160AU and 1.5° in longitude, and an initial guess at the density, $\mathbf{x}^{(0)}$, is

CME	Craft	d_i	Φ_{IP}	t_{insitu}	V_{sheath}	N_{sheath}	B_{max}	B_z	Dst
1	ST-B	1.0280	21	29/04/08 14:10	430±11	16±6	14.0	-9.5	...
2	ST-B	1.0542	-11	06/06/08 15:35	403±16	15±7	14.6	-8.6	...
3	ST-B	1.0548	19	07/06/08 12:07	384±17	13±7	12.5	-11.3	...
4	Wind	0.9840	-10	16/12/08 06:36	355±9	16±4	10.0	-7.6	...
5	STB	1.0263	34	31/12/08 01:45	447±10	7±3	9.5	-6.7	...
6	Wind	1.0023	28	18/02/09 10:00	350±8	22±4	12.3	-9.4	...
7	Wind	1.0004	19	05/04/10 07:58	735±18	10±2	21.5	-14.6	-81
8	Wind	1.0021	34	11/04/10 12:14	431±18	10±1	12.7	-8.6	-51
9	Wind	1.0132	-8	28/05/10 01:52	370±10	19±4	13.7	-12.9	-85
10	Wind	1.0161	-13	20/06/10 23:02	400±6	8±3	8.6	-2.8	...
11	ST-B	1.0604	-28	03/08/10 05:00	632±47	4±4	33.2	-30.2	...
12	Wind	1.0146	43	03/08/10 17:05	581±16	10±2	19.2	-11.2	-67
13	Wind	0.9881	40	18/02/11 00:48	497±27	25±11	31.8	-24.3	...
14	Wind	1.0145	25	04/08/11 21:18	413±12	6±1	10.1	-8.1	-107
15	Wind	1.0072	24	09/09/11 11:46	489±47	12±14	23.3	-21.4	-69
16	Wind	0.9946	15	24/10/11 17:38	503±15	26±4	24.3	-22.1	-132
17	Wind	0.9841	36	22/01/12 05:28	415±18	26±17	30.8	-27.9	-69
18	Wind	0.9844	33	24/01/12 14:36	638±34	8±2	30.5	-15.7	-73
19	Wind	0.9924	41	07/03/12 03:28	501±65	14±5	18.8	-18.2	-74
20	Wind	0.9927	42	08/03/12 10:24	679±44	12±4	30.4	-18.4	-131
21	Wind	0.9938	6	12/03/12 08:28	489±23	24±9	29.2	-23.6	-50
22	Wind	1.0055	20	23/04/12 02:14	383±8	24±7	15.9	-15.3	-108
23	Wind	1.0160	28	16/06/12 19:34	494±29	50±24	41.0	-21.0	-71
24	Wind	1.0165	21	14/07/12 17:38	617±39	16±6	27.7	-18.3	-127

Table 5.2: Table 2 from Möstl et al. (2014), which lists the Earth directed CMEs that have occurred during the STEREO mission. The CMEs to which the tomographic inversion has been initially applied are highlighted.

determined using the relation $n_e = 7.8r^{-1.94}$.

The elements of the regularisation matrices, $\mathbf{D}_r^2$ and $\mathbf{D}_\phi$, are determined using a finite difference scheme and the regularisation parameter, λ, is selected. Frazin (2000) presents a qualitative discussion of the choice of regularisation matrix but was inconclusive about which choice was most worthy of use. The value of λ used here is chosen such that the weighting of the regularisation matrix is approximately 5% of the observation matrix. This is found to ensure smoothness in the resulting density, without preventing the solving algorithm from converging toward the data.

The solving algorithm is then performed for 100 iterations or more, until the relative RMS residual between the data and solution is less than 10%. This second condition is found to be unnecessary, because the residual after 100 iterations is typically less than 5%. The results of the tomographic inversion for the 12 images between 00:09 and 07:29UT are shown in figures 5.13 and 5.16.

The electron density in the resulting maps has been multiplied by r^2 in order to remove the presence of the steep drop off from the plots. A clear peak in density associated with the CME can be observed originating near the Sun and propagating outward towards Earth. However, the morphology of the CME bears little resemblance to the accepted morphological model of a dense shell preceding a flux rope. Instead, the CME appears to be rather amorphous, which is due to both the regularisation parameters smoothing the resulting densities and by the fact that we are unable to reconstruct detailed morphology with only two points of view available.

The wide field of view provided by HI allows a technique for tracking CMEs developed by Sheeley et al. (1999) and adapted for the STEREO mission (Davies et al., 2009). Elongation-time plots, or *J-maps*, are created by stacking rows of pixels from successive HI images one after another. This is shown in the top and bottom row of figure 5.12, which serves to highlight moving structures. The left and right columns in this figure correspond to ecliptic data from HI-1A and HI-1B, respectively, during a 48 hour period ($12 - 13^{th}$ December 2008). The top row shows observed radiance (L2 data, with the star-field removed), whilst the bottom row shows simulated radiance produced by applying (5.16) to the densities obtained from the tomographic inversion. The middle row shows the fractional residual between the two. Specifically, the simulated radiance is the radiance we would observe with HI if our estimate of density were present in the ecliptic plane. Comparison of the observed and simulated intensities therefore provides a means of testing the consistency between the tomographic reconstruction of n_e and the HI image data. The RMS value of the residual over all pixels is typically a few percent over all the images, showing good agreement with the data, however certain regions exist where the estimate is less successful. The two vertical structures observed in HI-1B at approximately 0100 and 0800UT on the 12^{th} are seen to produce large residuals in both the HI-1A and HI-1B solutions at the corresponding times. This is due to erroneous background subtraction in the HI-1B images caused by sudden small movements experienced by the instrument. As

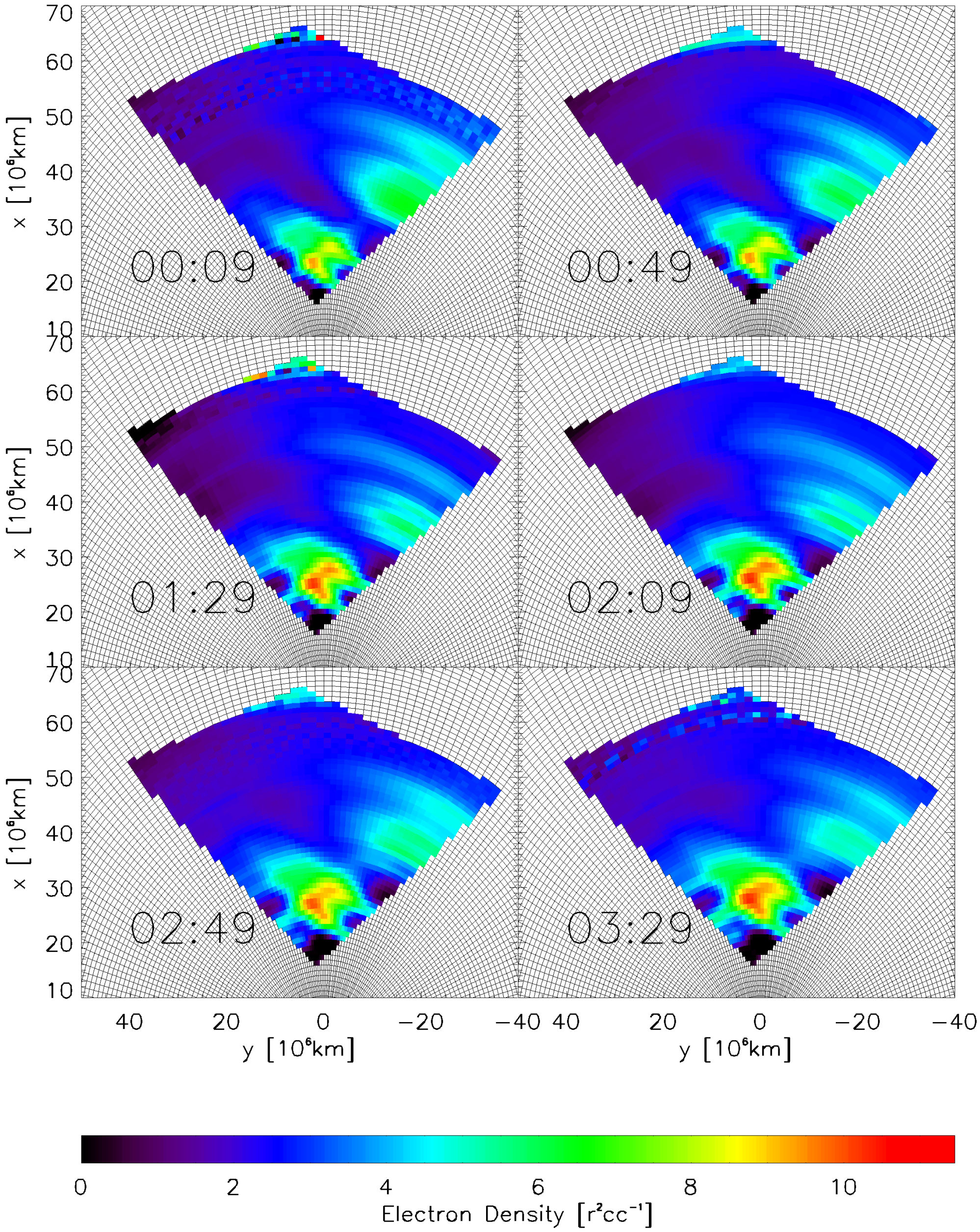


Figure 5.10: Resulting density maps from the tomographic inversion for the first six images on 13 December 2008. The results are displayed in the Heliocentric Earth Ecliptic (HEE) coordinate system, where the Sun is at the origin and the Earth lies at 1AU on the x axis. The resulting densities are multiplied by r^2 in order to remove the radial drop off.

a result, it may be expected that the density estimates based on these images to be poor. The residual values have a tendency to increase at larger elongation angles, particularly in the HI-1A observations above 20°. This may be reduced by increasing the significance

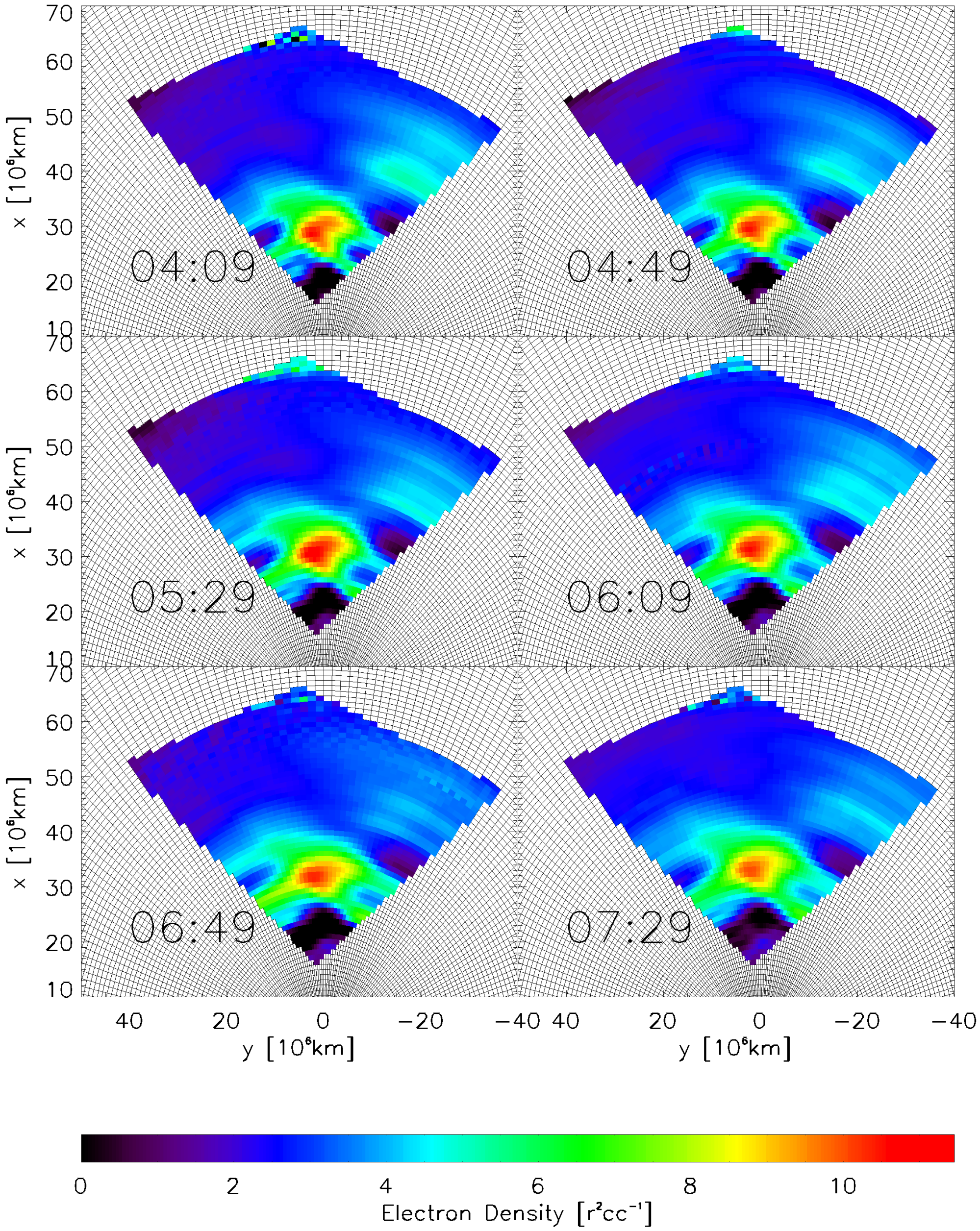


Figure 5.11: The same format as figure 5.13 produced from the following six images.

of the weighting factor, which was introduced specifically for this purpose. However, the CME is observed more clearly below 20°, indicated by the over-plotted diamonds, where the simulated radiance is in good agreement with the data. Some of the differences between the observed and reconstructed intensities are accounted for by the regularisation, which prevents the solving algorithm from converging to satisfy the data. The position of the peak in CME density, indicated by the diamonds, can be seen to follow the along a bright front-edge in each of the J-maps, which suggests that the solution has been successful in recreating the position of the CME within the ecliptic.

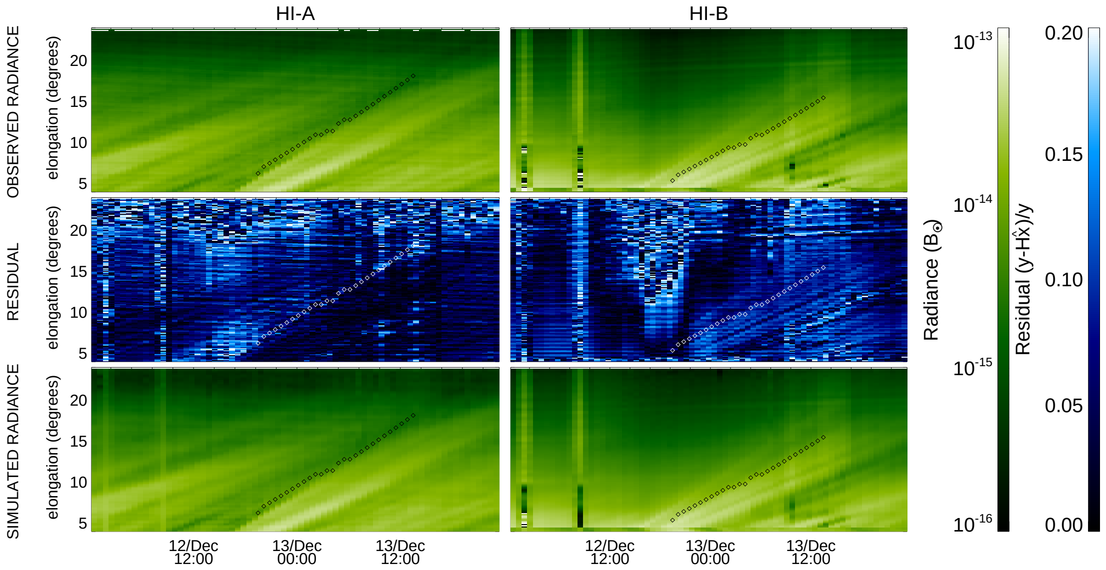


Figure 5.12: Observed and simulated intensities for 12-13 December 2008, where the left and right hand columns represent STEREO-A and -B respectively. The top row of plots shows J-maps produced from background subtracted, ecliptic intensities, whilst the bottom row show reconstructed intensities from the densities calculated for this CME. The middle row of plots shows the relative residual on a scale of 0 to 20%. The diamonds over-plotted in each panel show the position of the density peak in successive images (figures 5.13 and 5.16).

Figure 5.13 shows the reconstructed electron densities along the Sun-Earth line, as a function of time, in a similar fashion to the J-maps in figure 5.12. However, the vertical axis shows distance, rather than elongation. The diamonds show the approximate position of the peak in density, which can be seen to propagate away from the Sun, appearing at 1900UT on 12^{th}. The vertical features present in the HI-1B plot can be seen to have caused the algorithm to arrive at physically inconsistent densities, however these features are confined to the times at which the two “bad” images occurred. Again the CME appears somewhat amorphous, as the density enhancement does not form a narrow shock, but rather is spread over a region of ~ 0.1AU. We do however see a region of low density behind the CME, as is consistent with established CME models and observations. The CME remains as a clear enhancement above the background density out to approximately 0.3AU, at which point it becomes difficult to distinguish from the background. This is because the CME observed in HI becomes very faint as it travels through the HI FOVs

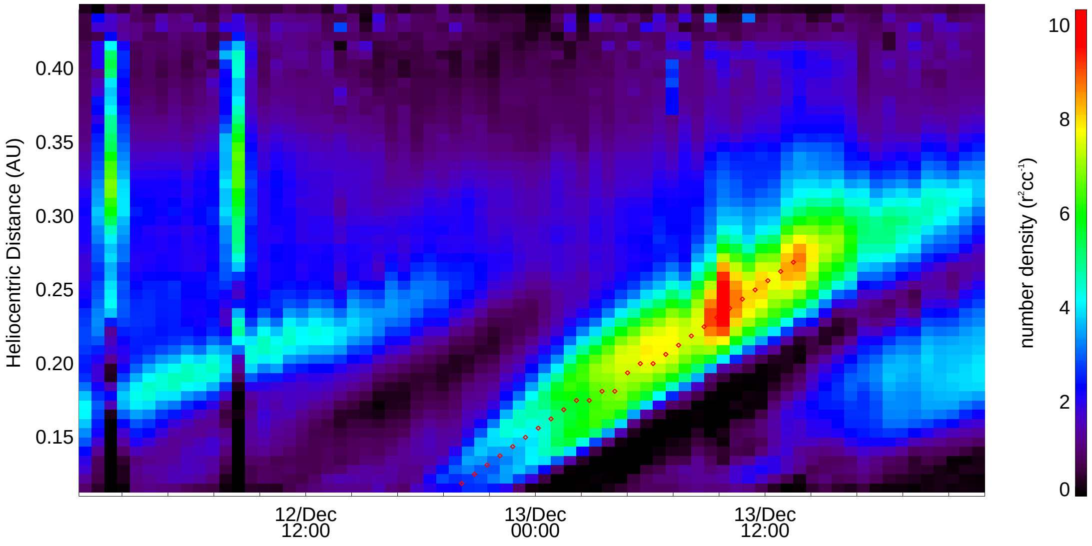


Figure 5.13: Electron densities resulting from tomographic inversion of HI data, as a function of time and distance from the Sun. The CME is the region of high density, which enters the FOV at approximately 1900UT on December 12^{th} and propagates outward. The red diamonds over-plotted represent the approximate position of the density peak at each successive time-step.

and cannot be clearly seen beyond approximately 20° elongation.

Whilst the consistency of the reconstruction with data has been demonstrated, a comparison with observations, such as density and velocity, are needed to justify the result. There exists a lack of these measurements within 1AU of the Sun, because they are conventionally measured in situ. With the exceptions of the MESSENGER and Venus Express spacecraft, these measurements do not exist within the orbit of Earth. The positions of these spacecraft, however, mean that they are of little use at this time (e.g. figure 5.5) because they do not lie within the region of space we are observing. An alternative is instead to extrapolate these tomographic results to 1AU and compare with near-Earth spacecraft. Because the CME in question is Earth-impacting, it is possible to make certain assumptions about its radial evolution in order to estimate how it will appear when it passes over the ACE and Wind spacecraft at L1. Many studies have been performed since the launch of STEREO, and prior to it, which address this problem.

Richardson et al. (2005) and Gazis et al. (2006) performed similar studies using large numbers of in situ measurements of CMEs, in order to estimate how their properties evolve, on average, with heliocentric distance. They determined respective power laws of $n_e(r) = 6.2r^{-2.3}$ and $n_e(r) = 5.74 \pm 0.27r^{-2.21\pm0.03}$ to describe the electron density as a function of radius, suggesting it is reasonable to assume the CME density follows a power law as it expands into the heliosphere. This result is in agreement with a theoretical study by Cargill (2004), who estimated the evolution of the density within a single CME, based on the Lorentz, gravitational and drag forces it experiences. The author's resulting equation describes a density that follows a power law with an unspecified index.

The Richardson et al. (2005) and Gazis et al. (2006) studies both conclude that, on average, CME velocity remains constant between the Sun and the Earth. Their innermost measurements come from the Helios spacecraft and so these conclusions are not applicable within 0.3AU of the Sun. McGregor et al. (2011) used the WSA-ENLIL model to demonstrate how the solar wind can undergo large accelerations within a few tenths of an AU, however, this is outside the region covered in the tomographic inversion in figures 5.13 and 5.16. This acceleration was found to be most significant for very fast and very slow solar wind, resulting from the interaction between these and the ambient solar wind. This conclusion is supported by the Cargill (2004) study and by observations from LASCO (Tappin, 2006), where changes in CME speed are shown to result from interaction with the solar wind, i.e. the difference in their respective speeds. For the two CMEs studied in this chapter, the velocities have been estimated to be close to the average solar wind speed of $\sim 450\text{kms}^{-1}$, for example by Möstl et al. (2014), so it is reasonable to assume that their velocities will not change greatly with r.

Lastly, assumptions may be made about the propagation direction of CMEs. Theoretical and observation modelling studies of CMEs (Cargill, 2004; Tappin, 2006; Tappin and Howard, 2009b; Wood et al., 2009) unanimously assume a structure that expands radially away from the Sun. That is, the angle of propagation does not vary with distance. Pizzo (1980) finds, based on an MHD model, that interactions with CIRs can alter the propagation direction of a CME due to the pressure gradient they impose upon it. This is, however, found to be a small effect compared with the outward expansion of the CME. The angle of propagation can therefore be assumed to be approximately constant.

By applying these considerations to the density estimates from STEREO, assumptions may be made about how the solar wind will appear at Earth. Along the Sun-Earth line, the peak density is plotted as a function of r and a power law fitted to the result. The radial position of the peak is plotted as a function of time and a linear fit used to find the speed. Results of these fits are shown in figure 5.14.

The resulting density estimate shows a poor agreement with the value measured in

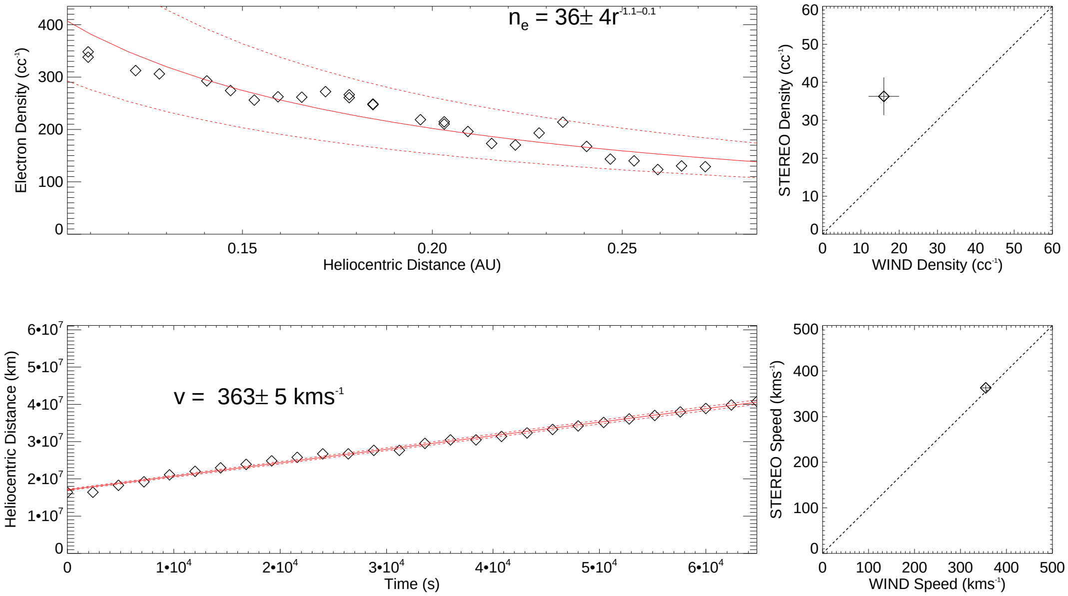


Figure 5.14: Evolution of CME parameters, with appropriate fitted trends for the December 2008 CME. The right hand plot in each case represents a comparison between the STEREO estimate and the value observed by Wind at L1, where the dotted line represents exact agreement. The top- left plot shows the peak electron number density along the Sun-Earth line, as a function of heliocentric distance, with a power law fitted to the data. The bottom left plot shows heliocentric distance as a function of time, along the Sun-Earth line. A linear fit has been applied to the data, the gradient of which assumes constant velocity.

situ, however the CME velocity is in much better agreement with that observed by WIND. These resulting profiles suggest that the assumptions made about the radial evolution of CME velocity are justified, however the density may be more difficult to predict. The tomographic inversion is quite limited in its ability to recreate CME morphology from just two points of view and so it may be that the density estimates are in fact quite inaccurate. Triangulating the position of a CME, and therefore determining its speed, are easier and so we may expect a good prediction of CME velocity as measured by WIND.

5.4.3 October 2011 CME

The tomographic reconstruction was applied to the CME launched on 22 October 2011, which is also known to be Earth-impacting. The technique has been carried out using the same method as was described for the previous event. The same parameters are used in the image processing and the same method is used to apply a correction to the data. The observed and simulated J-maps are shown in figure 5.15 and the reconstructed densities are shown in figure 5.16 over a 48 hour period ($22 - 23^{rd}$ October 2011).

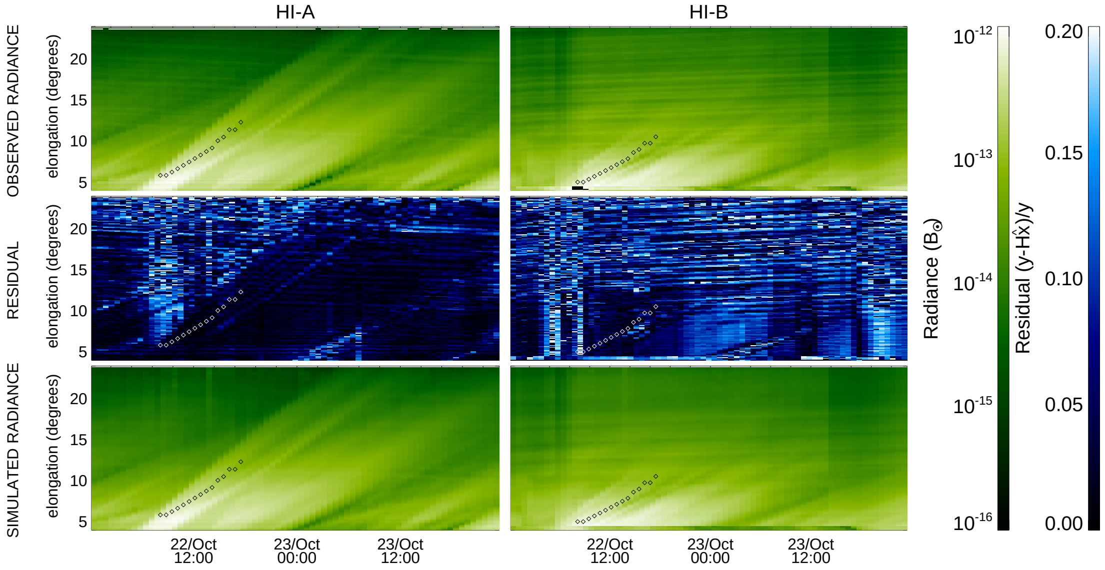


Figure 5.15: The same format as figure 5.12, but for the October 2011 CME.

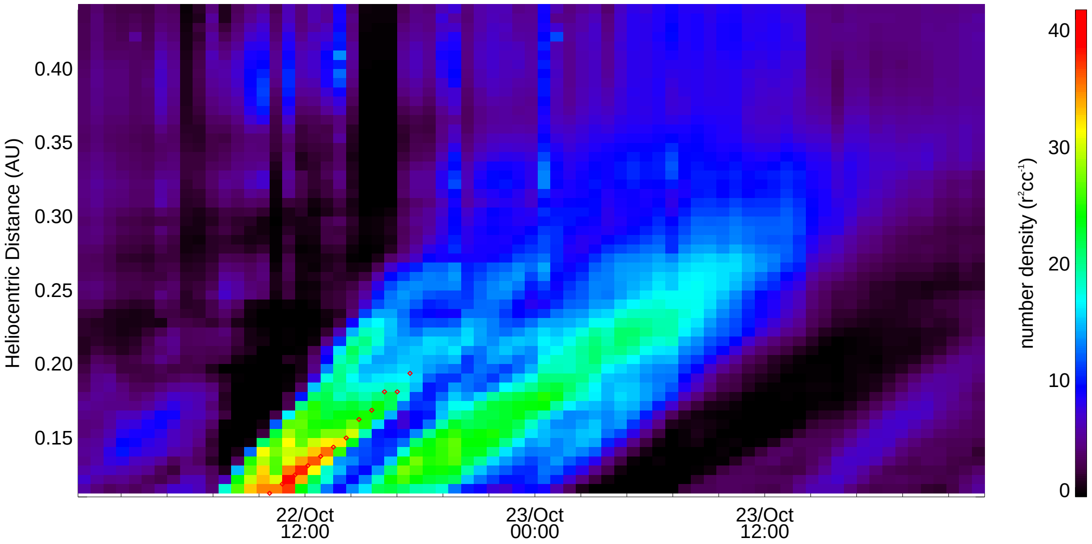


Figure 5.16: The same format as figure 5.13, but for the October 2011 CME.

The observed and simulated radiances in figure 5.15 show a reasonable agreement, however the densities determined during this period appear to represent the data less well than for the previous CME. The HI-1A estimate shows quite small residual values but those for HI-1B are typically between 10 and 20%. This is due to a greater density of background stars observed by HI-1B during this interval because the instrument is pointing close to the galactic plane. The median filtering technique applied to remove the stars is ineffective under these circumstances and causes starlight to remain in the image. Figure 5.17 is included to illustrate this effect. It shows the observed radiances from each of HI-1A and -1B at 0009UT on 23^{rd} October. It can be seen that the HI-1B data appears less smooth as a result of the background starlight. The regularisation included in the solution forces the density to have a smooth gradient, which results in worse agreement between the solution and the data. This is observed as a larger residual in figure 5.17 and the HI-1B plot in figure 5.15.

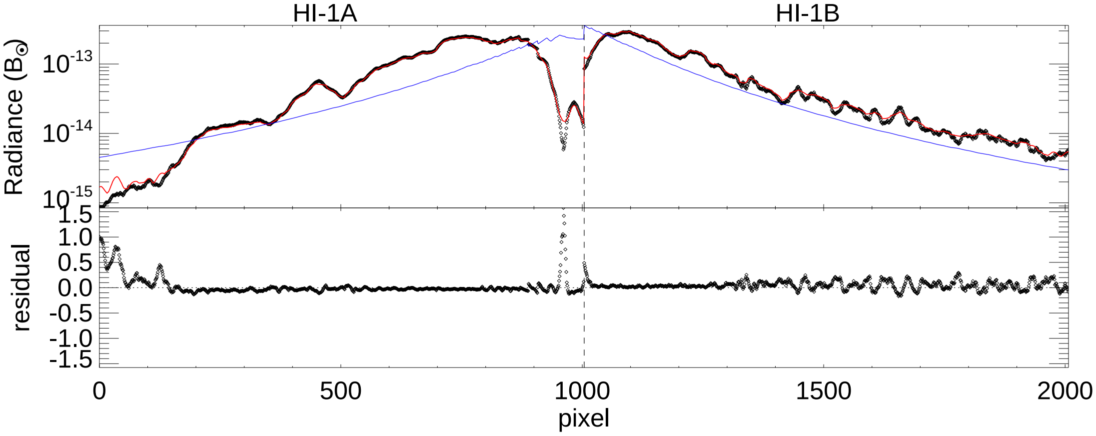


Figure 5.17: (Top) Observed (black) and reconstructed (red) radiances at 0009UT on 23^{rd} October 2011 from HI-1A and -B in the ecliptic. The blue line represents the radiance that would be expected from a density of $7.8r^{-1.94}$ electrons cc^{-1}, which is used as the initial guess in the solution. (Bottom) Fractional residual between the data and the solution.

As with the December 2008 CME, figure 5.15 shows a period of erroneous HI-1B data, this time between 0600UT and 1000UT on 22^{nd}, which is around the time the CME enters the fields of view. This results in larger residual values for both the HI-1A and HI-1B solutions. The HI-1A results show a greater discrepancy with the data above 20°, as was seen with the previous CME, whilst the HI-1B results are quite poor above 12° due to the greater starlight contribution. This means the CME cannot be tracked much beyond this point. The diamonds over-plotted, representing the peak density, appear to lag behind the bright fronts observed in both the HI-1A and -B observations. This may be caused by projection effects resulting from the increased spacecraft separation[5] by this point in the mission. The CME is an extended object, whilst the diamonds in figure 5.15 represent a single point within it. The observed bright edge in each of the J-maps may in fact be quite

---

[5] STEREO-A and -B were at respective HEE longitudes of $-100.3°$ and $+105.3°$ on 23^{rd} October 2011. Their positions on 13^{th} December 2008 were $-44.5°$ and $+42.3°$.

separate from the densest part of the CME.

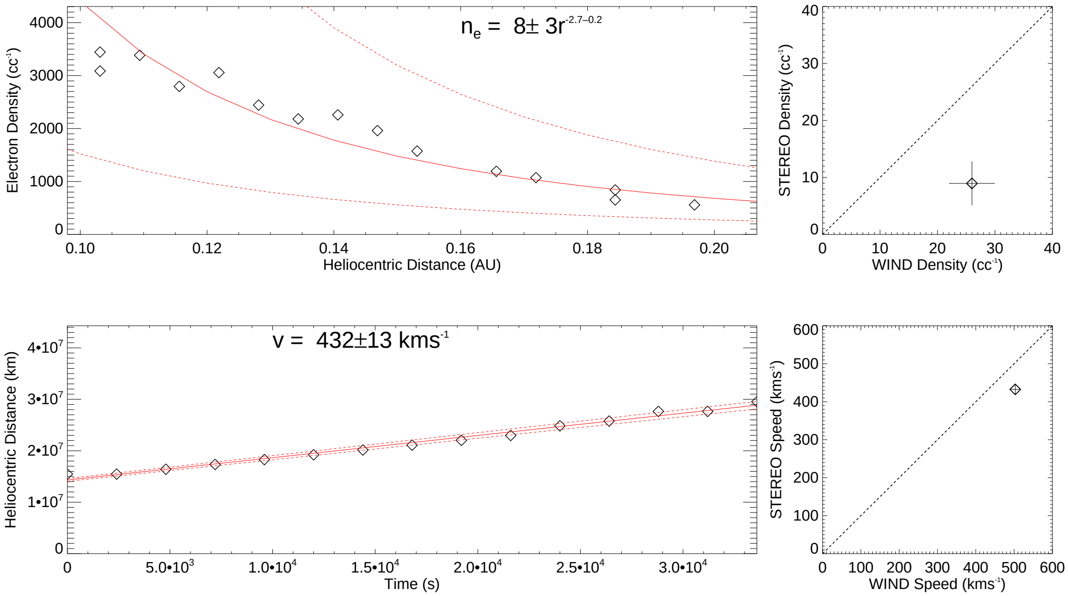


Figure 5.18: The same format as figure 5.14, but for the October 2011 CME.

Finally, figure 5.18 shows the density and velocity profile for this Earth-directed CME, based on the reconstructed density. As with the December 2008 CME, it is seen that the velocity is in good agreement with that observed by WIND, however the density profile is again a poor estimate. This corroborates the previous outcome that the method is not very accurate at estimating CME densities themselves. However, as a means to triangulate the position of these density enhancements and track them through the heliosphere the method proves quite successful.

Summary

This chapter demonstrates the effectiveness of remote sensing measurements in estimating CME properties, beyond the $\lesssim 10\mathrm{R}_\odot$ that the method has typically been applied to in previous studies. The tomographic inversion of white light images from the HI-1 on board each STEREO spacecraft has been demonstrated in its ability to reconstruct the density of two Earth-directed CMEs. It was possible to follow the December 2008 CME out to beyond 0.3AU, or nearly $70R_\odot$. The second event was more difficult to track, due to starlight observed by HI-1B, however the CME was tracked to 0.2AU, or over $40R_\odot$. In both cases, it was possible to make a good prediction of the speed of the event as was measured by the WIND spacecraft. This would in turn mean that the technique has the potential to be used to make CME arrival time predictions, if the propagation direction could be determined by some other means (eg. the methods of Davies et al. (2012); Davies et al. (2013)). The technique was successful in producing CME density estimates that were consistent with the data, however they were less than successful at predicting the densities that were observed by WIND. This is due to the fact that two points of view are insufficient to reproduce detailed structure. Whilst a region of enhanced density can be pinpointed, the distribution of density within that region is more difficult to determine. A disappointing aspect of these results is that there were only two CMEs identified from the list in table 5.2 to which the technique could be applied. A more recent and much more thorough list of nearly 2000 HI CMEs has been produced as part of the EU Heliospheric Cataloguing, Analysis and Techniques Service (HELCATS) project, which contains a subset of 393 events that are observed by both STEREO spacecraft. This includes potentially many more than two CMEs on with the tomographic inversion may be tested.

Algorithms used in tomography, when applied to CMEs observed in coronagraph studies, are generally accompanied by some forward modelling of the CME position and morphology (eg. Frazin et al. (2009); Frazin (2012)). This is because the two or three point of view available are too few to determine detailed structure. Similar forward modelling techniques would need to be included in the HI-based tomography in order to overcome its weakness in predicting densities. Again the HELCATS project contains a large number of CMEs, to which various forward and inverse modelling techniques are being applied. These models would serve as an excellent starting point for trying to improve the tomography.

Davila (1994) used simulated data to study the effectiveness of heliospheric tomography as a function of spacecraft number and separation. Unsurprisingly, the agreement between the reconstruction and simulation improved greatly as the number of spacecraft increased from two. Additionally a separation angle close to 180° was shown to give a poor result, which has imposed a large restriction on the periods in which HI data may be used for tomography. Potential future mission setups that would address these issues would be to position two spacecraft at the L4 and L5 points, where the spacecraft separation is close to optimal. Any increase in the overall number of spacecraft would improve the results greatly. Both of these mission concepts were in fact put forward for the original STEREO mission, however, due to compromises with other science requirements, the current setup was selected. Whilst these suggestions may seem a little optimistic, several scheduled, and proposed, missions will be carrying remote sensing instruments in to space in the coming years. Solar Probe Plus and Solar Orbiter are two missions with heliospheric imaging capabilities that are both scheduled for launch in 2018. Further to this, proposed future instruments include polarised heliospheric imagers, which would improve tomography by allowing polarimetric reconstruction of density structures.

The main restriction on performing tomography at large elongation angles is that the Thomson scattered brightness falls off steeply with radius and is orders of magnitude lower than the F-corona and background star-field. These issues are addressed, to some extent, using the techniques described in this chapter, but more thorough image processing is required for tomography to be performed effectively at large elongation angles. If this could be achieved, then it potentially provides a means of directly measuring the geoeffective properties of CMEs in transit, as they travel towards the Earth, beyond the few solar radii to which these techniques are currently limited.

This page was intentionally left blank

Chapter 6

The Effects of Electric Field Variability on the Neutral Atmosphere

The subject of this chapter is the electric field input in CMAT2. The chapter begins with a summary of the K_p dependent electric field model of Foster et al. (1986), which was used to drive CMAT2 prior to the work completed in this thesis. Following this is a description of how newly available SuperDARN data (section 3.2.1) are used to replace this model, using alternative measurements of electrostatic potential. The remainder of this chapter then focuses on a comparison between these two models, specifically the limitations of Foster's model in its ability to be used in driving a GCM.

For the purposes of ionosphere-thermosphere modelling, it is necessary to quantify geomagnetic activity in a way that can be used to drive the model's ionosphere and to derive the resulting energy and momentum input to the neutral atmosphere. In the unmodified version of CMAT2, this is achieved by using proxies to determine the three heating rates that result from absorption of solar radiation, particle precipitation and Joule heating. Firstly, the F10.7 index is used to select a value for solar irradiance per unit wavelength, $I_\infty(\lambda)$. This acts as an input to (2.73), from which the heating rate $Q(\lambda)$ is determined in (4.57), which in turn contributes to the energy equation, (4.44). The K_p proxy is used to determine both the particle precipitation and Joule heating. For a given K_p there is a corresponding particle precipitation pattern in the form of the total ionisation rate, Q_T, as a function of position over the CMAT2 grid. This determines the ionisation rates for each of the major species from (4.65), (4.66) and (4.67). Finally, each K_p also has a corresponding potential distribution, Φ, which is used to find the electric field, $\mathbf{E}$. This contributes to Joule heating via the energy equation (4.44) and neutral velocities via ion-neutral collisions in the momentum equation (4.42) and (4.43).

Using SuperDARN data instead of Foster's model removes the dependence of the electric field on the K_p index. The SuperDARN data used in this study cover the years 2000 to 2010. The data are available every two minutes, containing the 49 coefficients A_{lm} and B_{lm} to 6^{th} order. CMAT2 operates with a 60s time step, so the process of implementing SuperDARN data into the model is then to call these 49 coefficients every 2^{nd} time step, along with the HMB, Λ_0, and to determine the electrostatic potential in each grid cell using (3.3) and (3.4). The conversion between the magnetic coordinate system of SuperDARN and the geographic coordinate system of CMAT2 is performed using the transformations presented in section 4.5. The model then uses a finite difference approximation between neighbouring grid cells to calculate the electric field vector from equation (3.1). Figure 6.1 shows an example of the electrostatic potential and electric fields plotted over a grid of the same resolution as CMAT2 (2° latitude and 18° longitude). The plots in the figure cover a half hour period, which illustrate the temporal variability within the data, particularly when compared with the Foster model that changes only on timescales of hours when the K_p value is updated. It is also apparent that the SuperDARN potential maps posses much greater spatial variability, whereas the Foster model generally maintains the same shape between different activity levels. This results from the way the Foster model is based on averaged data, rather than the (approximately) instantaneous measurements of SuperDARN.

There are two main points that are to be addressed in this chapter with regards to the ability of these models to produce realistic energy and momentum in CMAT2. Firstly, section 6.1 presents a study of the abilities of the K_p index and IMF as proxies for predicting the ionospheric electric field when compared to SuperDARN observations. Section 6.2 focuses on the electric field variability within these models and how this effects the energy and momentum input to the CMAT2 neutral atmosphere.

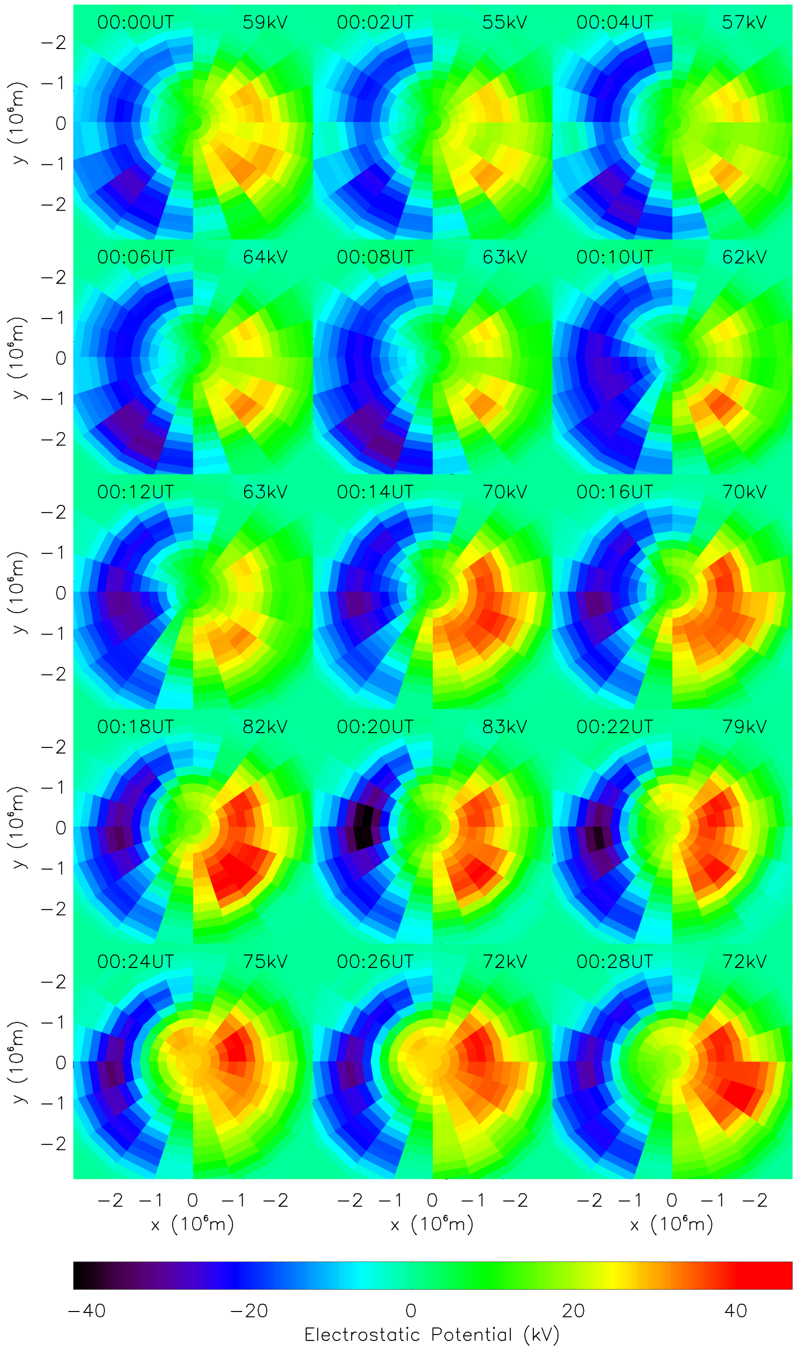


Figure 6.1: Time series of SuperDARN electrostatic potential plotted in magnetic coordinates with a resolution of 2° magnetic latitude and 18° magnetic longitude, equal to the resolution of CMAT2 in geographic coordinates. The series covers a half hour period between 0000 to 0030 UT on 20 March 2001, with a time resolution of 2 minutes and proceeds from left to right and top to bottom, with 12 MLT to the top of the plot.

6.1 The K_p Index and IMF as a Proxy for Ionospheric Electric Fields

The K_p index (Bartels et al., 1939) represents a measurement of the maximum deviation in magnetometer readings over a three hour period from ground stations distributed in longitude and latitude. The unmodified version of CMAT2 uses the electric fields derived by Foster et al. (1986), in which observations from the Millstone Hill radar are bin-averaged according to the K_p index. Based on either a user-selected or a historic value of the K_p index, CMAT2 then calls the corresponding electric field according to Foster's model. A more sophisticated alternative to Foster's model is that of Weimer (2005), which is driven using the IMF vector, $|\mathbf{B}|$, the solar wind speed, $|\mathbf{v}|$ and electron density. Whilst Weimer's electric field has yet to be successfully implemented into CMAT2, the relationship between Foster's model, Weimer's model and SuperDARN are compared here.

This section compares the output of Foster and Weimer's model to that of SuperDARN as a means to assess the ability of K_p and the IMF vector as proxies for electric fields. The SuperDARN data used here are restricted to the years 2000-2003, during the maximum of solar cycle 23. The SuperDARN data available for use in CMAT2 a further seven years to the end of 2010, however after it is found that little back-scatter occurs during this period. As a result to SuperDARN data become more reliant on the empirical model, which we wish to avoid. As an extra measure to avoid this, a further restriction is placed on which periods of data are used. In a given hemisphere, we discard any time-steps where fewer than 200 data points contribute to the potential pattern.

6.1.1 The K_p Index

SuperDARN data are selected that satisfy the above criteria. For each potential pattern, the cross polar cap potential difference, Φ_{pc}, is calculated and binned according to the corresponding K_p index between 1+ and 6−, which is the range used by CMAT2. This represents the difference between the maximuma and minimum electrostatic potential values measured in a given hemisphere and is therefore a good proxy for geomagnetic activity. The value of Φ_{pc} in each bin is also calculated from the corresponding Foster fields. Figure 6.2 shows the resulting potentials as a function of K_p index. The cross-cap potential from Foster is dependent on K_p and therefore the plot corresponds to the fixed values in table 4.1, increasing from 25.78kV at $K_p = 1+$ to 79.87kV for $K_p = 5-$. Values in the $K_p \geq 6-$

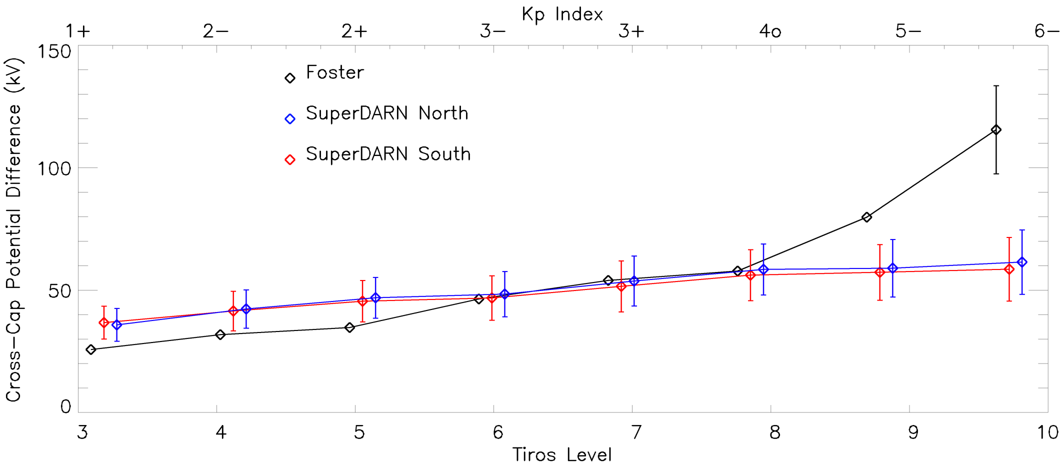


Figure 6.2: The average potential drop, Φ_{pc}, across the polar cap binned according to Kp, or equivalent TIROS, level for the duration of 2000-2003. The black line indicates Φ_{pc} from Foster's electric field model. The blue and red lines correspond to values measured by SuperDARN in the northern and southern hemisphere, respectively. The error bars in each case corresponds to the standard deviation in each bin.

bin are calculated from (4.62) and (4.63), an therefore vary depending on the particle precipitation energy. The SuperDARN cross-cap potentials, in each hemisphere also show an increase for each successive K_p value, although the range is much smaller than for Foster's model. The SuperDARN cross-cap potential increases from 36 and 37kV to 61 and 59kV in the northern and southern hemispheres respectively. This may seem counter-intuitive because the K_p index is a measure of geomagnetic activity, which is known to be associated with an expansion of the auroral oval and an increase in Φ_{pc}. However the K_p index is a measure of the maximum deviation in magnetometer readings over a three-hour period. Within each three-hour period, there may be up to 90 SuperDARN potential maps that contribute to a given bin in figure 6.2. Of these SuperDARN potential maps, we may expect the maximum Φ_{pc} to correspond to the K_p index, however the rest may be significantly lower. The consequence of this is not that the SuperDARN Φ_{pc} does not vary according to geomagnetic activity, but that the K_p index is a poor proxy to drive a GCM. This was not the purpose it was devised for and instead is intended to be a measure of the maximum severity of geomagnetic activity within a given period.

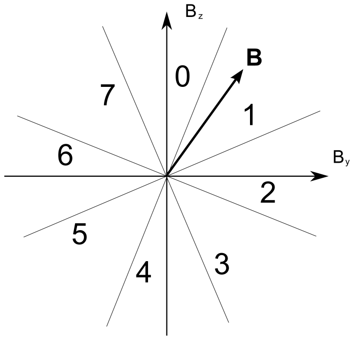


Figure 6.3: IMF sector as a function of magnetic field vector.

6.1.2 The Interplanetary Magnetic Field

Magnitude of the Electric Field

According to Weimer's model (Weimer, 2005) and the energy coupling function (Perreault and Akasofu, 1978; Kan and Lee, 1979; Akasofu and Ahn, 1980) the cross-cap potential is strongly dependent on the IMF and solar wind parameters, $\mathbf{B}$, $\mathbf{v}$ and n_e. The most significant of these being the direction of $\mathbf{B}$, i.e. the IMF sector, which corresponds to the region in which the vector $\mathbf{B}$ lies (figure 6.3). Accordingly, the same years (2000-2003) of K_p values have been determine the Foster electric fields, which are then binned by IMF sector. The same SuperDARN data from this period are also binned by IMF sector. IMF values are taken from NASA's High Resolution OMNI data-set. Finally the cross-cap potentials for Weimer's model based on average solar wind conditions ($|\mathbf{v}| = 450\text{kms}^{-1}$, $|\mathbf{B}| = 5\text{nT}$ and $n_e = 4\text{cc}^{-1}$) are binned in the same way.

Figure 6.4 shows each of the cross-cap potential values plotted as a function of sector. Sector 0, in the centre, corresponds to strongly positive B_z values and sector 4, at each edge, corresponds to strongly negative B_z. The cross-cap potentials for Weimer represent average solar wind conditions ($|\mathbf{v}| = 450\text{kms}^{-1}$, $|\mathbf{B}| = 5\text{nT}$ and $n_e = 4\text{cc}^{-1}$) in each sector. These values are taken from Weimer (2005) where they are presented as a function of IMF sector. The cross-cap potential in the Weimer model increases from 24kV in sector 0 to 102kV in sector 4. When binned by IMF sector the Foster cross-cap potential shows no significant increase between negative and positive B_z values from 37kV in sector 0 to 46kV in sector 4 and is far smaller than the range seen in the Weimer model. The SuperDARN data follow a more similar trend to the Weimer model than to Foster's. The respective minimum values in the northern and southern hemispheres are 23 and 21kV, however the

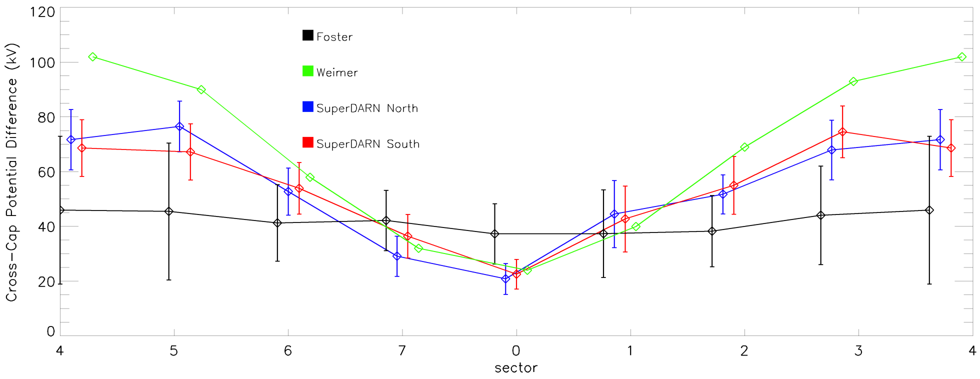


Figure 6.4: The cross-cap potential drop averaged according to IMF sector for the year 2001. The black line corresponds to Foster's model, the green to Weimer's model and blue and red represent SuperDARN data in the northern and southern hemisphere, respectively. The standard deviation in each bin is shown by the error bars.

respective maxima occur in sectors 5 (75kV) and sector 3 (77kV). This would suggest a small B_y dependence, however, the values from each hemisphere are within one standard deviation in each bin. The SuperDARN values are generally lower than the Weimer model and this effect is most strongly seen in sectors 3 to 5 when B_z is negative. The consequence of figures 6.2 and 6.4 is that the IMF sector and K_p index are very different ways of quantifying geomagnetic activity and bear little resemblance to each other.xxs

Extent of the Heppner-Maynard Boundary

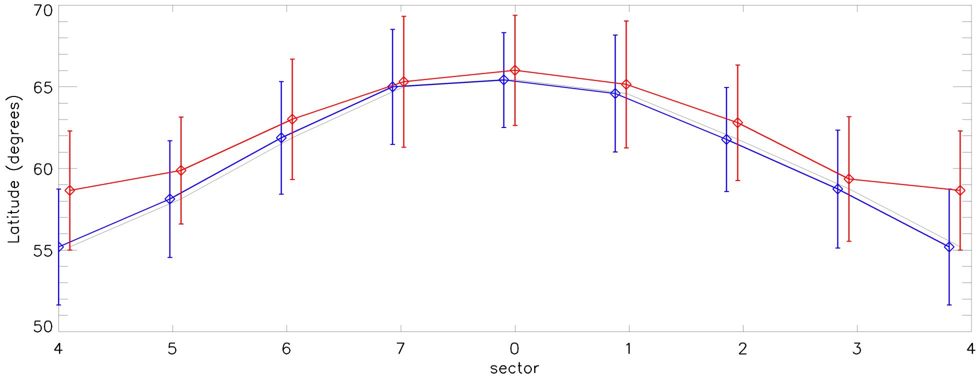


Figure 6.5: The absolute value of latitude of the HMB in both the northern (blue) and southern (red) hemisphere as a function of IMF sector. The dashed lines indicate the standard deviation in each bin.

A further property of the convection provided by the SuperDARN data sets is it's latitudinal extent, given by the HMB, Λ_0. Figure 6.5 shows the average position of the HMB, as a function of IMF sector, for the same four years 2000-03. Both hemispheres show an increase in the extent of the HMB, as B_z becomes increasingly negative. The boundary extends from approximately $|\phi| = 66°$ in sector for both hemispheres to 55° and 59° in sector 4 for the northern and southern hemispheres respectively.

6.1.3 Discussion

Figure 6.2 showed the Foster and SuperDARN cross-cap potentials as a function of K_p index. Foster's model, which uses this index as a proxy, shows an increase potential as the K_p index increases, whilst SuperDARN shows little correlation to it in either hemisphere. The K_p values most frequently measured correspond to the TIROS levels between 4 and 7 (Foster et al., 1986), where the Foster model gives a similar representation of the size of the SuperDARN electric field. However, at high K_p values, particularly above $6-$, the Foster model appears to be a poor estimate of the potential, when compared to SuperDARN. The potential produced by Foster is significantly larger under these conditions, suggesting that it would produce very different effects in a GCM; driving much greater wind speeds and heating.

When binned by IMF sector (figure 6.4) the Weimer and SuperDARN cross-cap potentials show a significant increase as B_z becomes strongly negative, whilst Foster's model shows little change for any IMF orientation. This implies that the IMF orientation, rather than the K_p index, is a better proxy to drive the electric field within CMAT2 if SuperDARN observations are considered to be accurate. This is because the K_p index represents a *maximum* value of magnetometer readings during a given three hour interval, which means the *average* of the 90 values measured by SuperDARN during the same period will necessarily be lower, assuming any peak in activity occurs on a time-scale of much less than three hours. This is seen in figure 6.2 for the three highest K_p values. Additionally, the amount by which the IMF can vary over a three hour interval (for example figure 6.6) is what causes the Foster field strength to be smoothed-out when binned according to sector in figure 6.4. While large K_p values are known to correlate with strong geomagnetic activity this does not mean K_p is a good proxy for the ionospheric electric field at a given instant and is therefore not suitable to drive a GCM. The IMF is a far better proxy because it is more directly connected to the state of the ionospheric electric fields at a given time.

It was noted in figure 6.4 that SuperDARN potential values are generally lower than those of the Weimer model, particularly for negative B_z. Cousins and Shepherd (2010) compare the results of various cross-cap potential measurements, including SuperDARN, Weimer's 2005 model based on DE2 observations and DMSP and incoherent scatter (IS) radar measurements. The authors show that the IS radars produce similar values to SuperDARN, however SuperDARN consistently underestimates the cross-cap potential when compared to the Weimer model and DMSP. This discrepancy is particularly large when B_z is strongly negative, as was also observed in figure 6.4. Gillies et al. (2011) propose that the cause of this may be that SuperDARN velocity measurements use an incorrect estimate of the refractive index in the regions of the ionosphere where back-scatter occurs. This underestimated velocity then produces an underestimate for the electric field strength or electrostatic potential. This would suggest that, whilst IMF orientation is a good measure of the how the ionospheric electric field varies, the absolute magnitudes seen in SuperDARN may be smaller than the true values present in the ionosphere. Use of SuperDARN as a convection model to drive CMAT2 will then be expected to produce underestimated Joule heating and wind speeds as a result of lower values of $|\mathbf{E}|$ in the energy and momentum equations. Possible corrections to the underestimated velocity measurements have been proposed. Gillies et al. (2012) demonstrate that taking two velocity measurements of the same scattering volume at different radar frequencies provide a means of measuring its refractive index. Spaleta et al. (2015) show that this may be improved with more thorough mathematical analysis of the dual-frequency measurements. These techniques, however, require changes to the way radar measurements are taken. They are therefore not applicable to existing measurements, such as the data used to drive CMAT2.

Figure 6.5 shows that the HMB, in both hemispheres, expands from $|\phi| \sim 65°$ for B_z strongly positive to around 10° further equator-ward for B_z strongly negative, when binned by IMF sector. CMAT2 has been modified to use the SuperDARN electric fields, however it is still reliant on the K_p driven TIROS model of particle precipitation, due to the absence of a matching precipitation model. The lack of correlation between IMF sector and K_p is significant because there are likely to be periods when the electric field and particle inputs do not match up spatially within the model. Because Joule heating in CMAT2 is dependent on both the conductivity and electric field strength, this may present difficulties when modelling heating accurately. Imber et al. (2013) show that the HMB, as measured by SuperDARN, may be used as a proxy for the location of the auroral oval. Whilst this suggests that there is potential to develop a complementary particle precipitation model, that task is beyond the scope of this thesis.

6.2 Electric Field Variability

Codrescu et al. (1995) state that GCMs based on averaged convection patterns, such as Foster and Weimer's models, will underestimate the contribution of Joule heating to the thermospheric energy budget. In these models the Joule heating is found to be proportional to the *square* of the *average* E-field, whilst the true Joule heating is proportional to the *average* of the *square* of the electric field. The significance of this is that the latter values will always be greater, resulting in an underestimate of Joule heating from averaged electric fields.

Codrescu et al. (2000) demonstrate this contribution of E-field variability to Joule heating by assuming the electric field values to follow a Gaussian distribution superimposed on the average e_m. That is, a given electric field value, e, is equal to $e_m + x$, where x is a Gaussian random variable with the following probability density function

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2\sigma^2}} \mathrm{d}x \tag{6.1}$$

The average, or expectation value, of the random variable x with probability density $f(x)$ is equal to

$$\overline{x} = \int_{-\infty}^{+\infty} x f(x) \mathrm{d}x \tag{6.2}$$

and the expectation value of e is obviously equal to e_m,

$$E(e) = \overline{e} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (e_m + x) e^{-\frac{x^2}{2\sigma^2}} \mathrm{d}x = e_m \tag{6.3}$$

However, the Joule heating, Q_j is proportional to the expectation value of $\overline{e^2}$

$$Q_j \propto \overline{e^2} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (e_m + x)^2 e^{-\frac{x^2}{2\sigma^2}} \mathrm{d}x = e_m^2 + \sigma^2 \tag{6.4}$$

The consequence of this is that both the mean electric field and the variability contribute equally to the Joule heating. When using averaged electric field models in GCMs, it is necessary also to quantify the variability in order to produce the correct physics.

This represents a further flaw with the Foster model, which is based on a three-hourly K_p index and contains no variability on this time-scale. The previous section demonstrated K_p to be a poor proxy for the electric field strength at a given moment, suggesting that Foster is unlikely to succeed at predicting either e_m or σ accurately. The Weimer model is instead based on geo-effective solar wind parameters, which are found to be a better proxy for electric field strength. Also, these parameters may be obtained from the OMNI data set with a time resolution of one minute, which would suggest that the Weimer model is also able to account for the variability, σ. Transmission between the magnetosphere and ionosphere is not expected to be perfect or instantaneous, so the variability seen in OMNI data does not translate perfectly to ionospheric variability. However an electric field driver with one-minute resolution, such as IMF, is likely to possess more variability than one driven with 180-minute resolution, such as K_p. Although the Weimer model is not implemented in CMAT2, the effect of variability within the model may be studied by comparing simulations using the Foster field to those using SuperDARN electric fields. Whilst Codrescu et al. (1995, 2000) address the effects of electric field variability on the Joule heating, similar reasoning may by applied to estimate the effect of variability on neutral wind speeds. At altitudes above the F-region ($\sim$ 240km), ion drag can be assumed to be the main force driving neutral wind speeds. It should then be expected that the mean neutral velocity is proportional to the mean of the electric field strength, independent of any variability,

$$V \propto \bar{e} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (e_m + x) e^{-\frac{x^2}{2\sigma^2}} \mathrm{d}x = e_m \tag{6.5}$$

These assumptions are tested by performing two CMAT2 simulations, each using different electric field models (Foster and SuperDARN), but with otherwise identical settings. The model runs are be performed over a five day period from March 16 − 20 2001, one using SuperDARN electric fields and one using Foster's model. The period is chosen to be during solar maximum because this is when the SuperDARN radars typically produce more data, and during equinox because the coverage is generally good in both hemispheres. During this specific five day interval SuperDARN was consistently observing back-scatter, minimising its dependence on the IMF-based empirical model. The SuperDARN-driven run uses northern hemisphere measurements to determine both the northern and southern electric fields. The purpose of which is to mimic the way in which the Foster fields behave for a fair comparison and to avoid any effects of hemispheric asymmetry, which are addressed in the following Chapter.

6.2.1 CMAT2 results

Figure 6.6 shows the state of the solar wind during the five day period simulation period, 16^{th} to 20^{th} March 2001, from the OMNI data set. The cross-cap potential during the same period is plotted in figure 6.7 from Foster's model, using the K_p values that were present. Over-plotted are the values from SuperDARN, in each hemisphere. Rather than plotting these at two-minute resolution they are averaged over three hourly intervals in an attempt to illustrate the relative values of e_m and σ. Also plotted is the value of Φ_{pc} determined from the energy coupling function. The increase in magnitude of Φ_{pc} in all three data sets in figure 6.7 at approximately 1200UT on 19^{th} March can be seen to coincide with an increase in solar wind activity in figure 6.6. There is a southward turning of $\mathbf{B}$, in addition to an increase in its magnitude and that of the solar wind speed and density. There is however, less of a response to these changes in the SuperDARN data than in the Foster model. At its highest value ($\sim$0700UT on 20^{th} March) the cross cap potential in Foster's model is more than double that of the SuperDARN measurements, whilst during quieter periods they remain approximately equal.

The relationship $\overline{Q_J} \propto e_m^2 + \sigma^2$ proposed by Codrescu et al. (2000) may be tested by comparing the output of each simulation. Assuming it is legitimate, Foster's model will produce Joule heating, $\overline{Q_f} \propto e_{mf}^2$, when averaged over a period of less than three hours[1]. SuperDARN will instead produce heating, $\overline{Q_s} \propto e_{ms}^2 + \sigma_s^2$, when averaged over the same period because it contains variability. If these relations are precise, then it would follow that the two simulations should produce relative heating rates according to $\overline{Q_f}/\overline{Q_s} = e_{mf}^2/(e_s^2 + \sigma_s^2)$. Likewise, the relative neutral wind speeds, averaged in the same way should obey $\overline{V_f}/\overline{V_s} = e_{mf}/e_{ms}$. To arrive at these averages, the values of $|\mathbf{E}|$, and Q_J in each hemisphere are averaged over one-hour periods and over the entire hemisphere, where Q_J is height integrated Joule heating. A similar process is applied to neutral speeds, $|\mathbf{V}|$, however only values at 400km altitude and absolute magnetic latitude ($|\phi_m|$) above

---

[1] In this context, e_{mf} refers to the average electric field value within CMAT2 from the Foster model. The subscripts e_{ms} and σ_s refer to the mean and standard deviation of the SuperDARN fields.

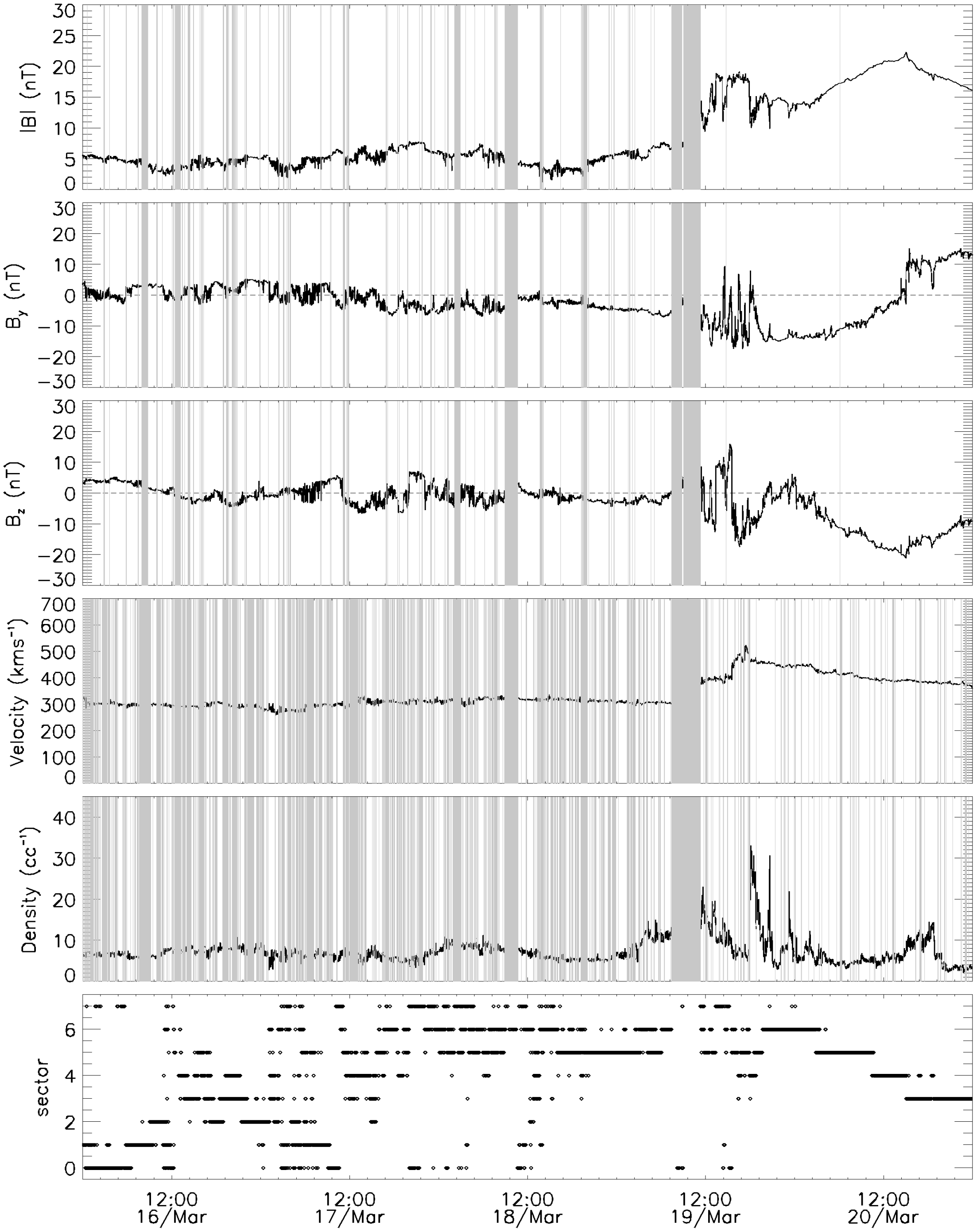


Figure 6.6: Solar wind and IMF parameters for 16 – 20 March 2001, from the OMNI data set. The top plot shows the magnitude of the IMF vector and the second and third plots show its y and z components. The fourth and fifth plots are the solar wind speed and density and the last plot indicates the IMF sector. The greyed-out areas indicate an absence of data.

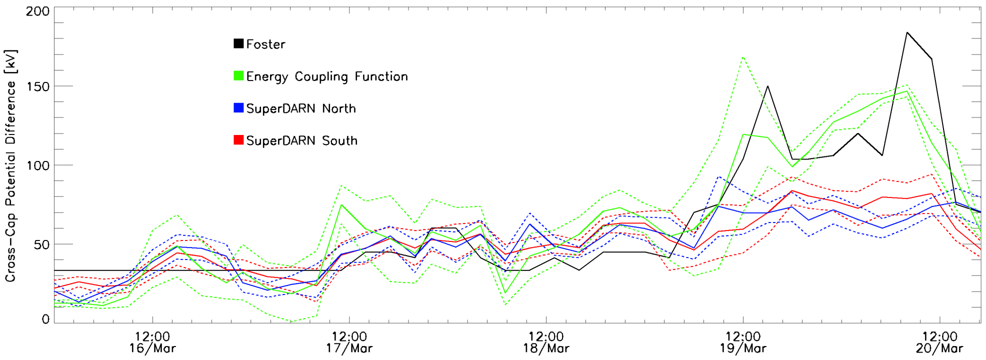


Figure 6.7: The cross-polar cap potential drop for 16 − 20 March 2001. The black line indicates the value corresponding to the Foster fields derived from the three-hourly K_p index. The green line indicates the mean of the value derived from the energy coupling function (2.79) using WIND data, binned every three hours. The respective blue and red lines correspond to SuperDARN data in the northern and southern hemispheres, binned every three hours. Dashed lines show the standard deviation in each case.

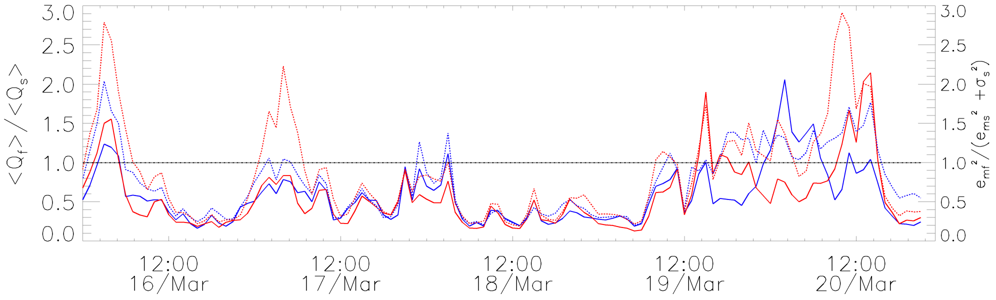
80° are included, where ion drag is most significant on the neutral atmosphere.

Effects on Joule Heating

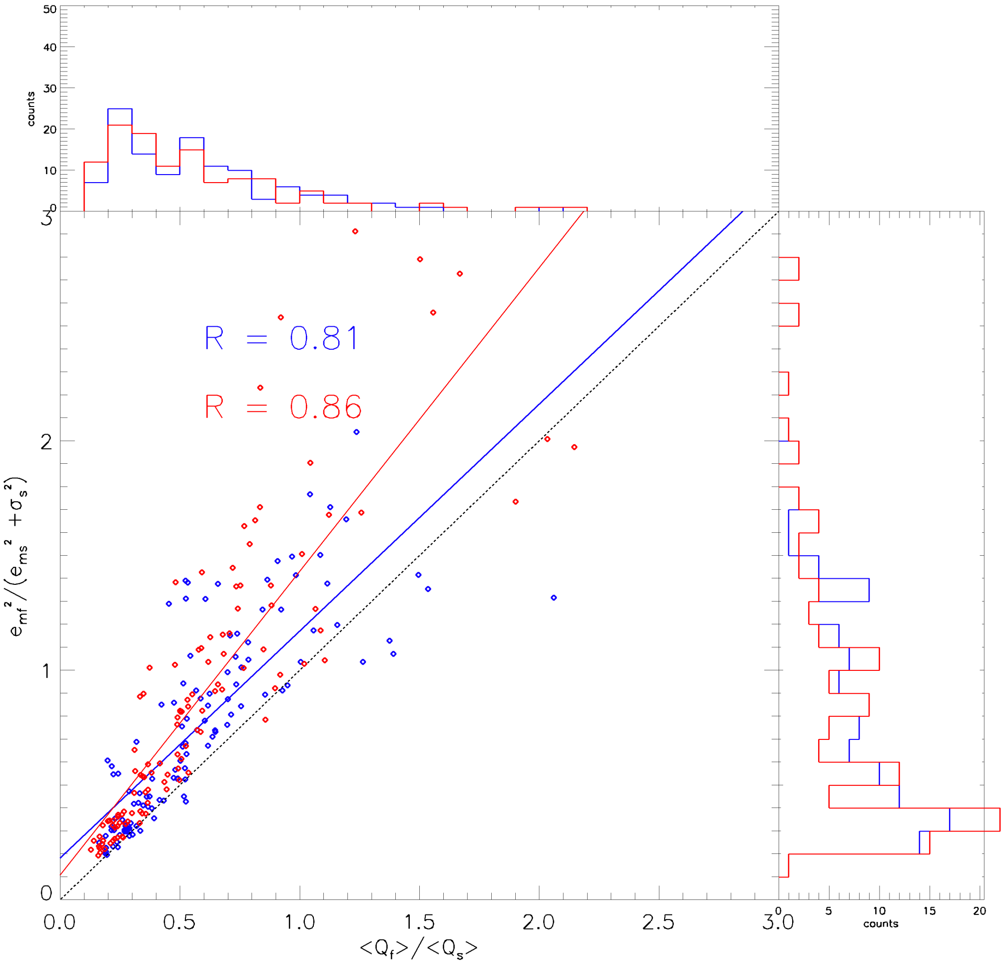
The ratio $\overline{Q_f}/\overline{Q_s}$ is plotted in figure 6.8(a), with $e^2_{mf}/(e^2_{ms}+\sigma_s)$ over-plotted. These two ratios are plotted directly against each other in 6.8(b). Figure 6.8(a) shows the heating in the SuperDARN driven model is consistently greater than that in the Foster model, for the period up to around 1200UT March 19^{th}. After this point, when the solar wind conditions become stronger, the Foster model produces more heating. Figure 6.8(b) shows a linear relationship between the two ratios, with respective R values of 0.81 and 0.86 in the northern and southern hemispheres. The regression lines have gradients of 0.99 ± 0.07 (north) and 1.32 ± 0.07 (south). This suggests that the assumed relationship between electric fields and heating is satisfied.

Effects on Neutral Wind Speeds

Figure 6.9 shows a plot in the same format but for the mean neutral wind ratio, $\overline{V_f}/\overline{V_s}$, and the mean electric field ratio e_{mf}/e_{ms}. Figure 6.9(a) shows the neutral speeds in the SuperDARN driven simulation to be typically greater than those in the Foster simulation, up to a factor of ~ 2, during the period up to 1200UT on March 19^{th}. This difference disappears when the solar wind conditions become stronger during the final 24 hours of the simulation. The scatter plot in figure 6.9(b) again shows a good agreement between the

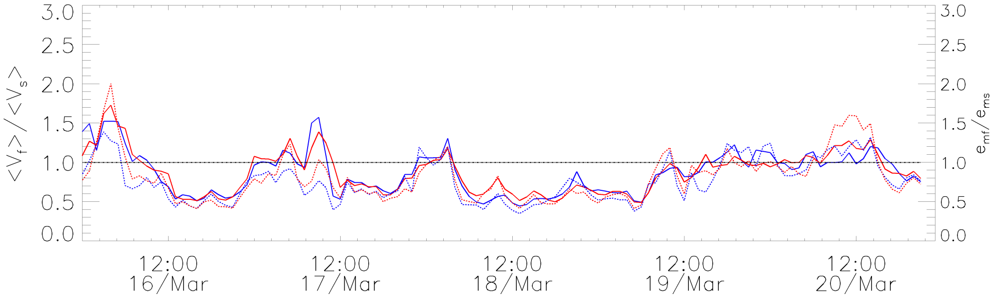


(a) The solid lines indicates ratio of mean Joule heating from the model run using Foster to the model run using SuperDARN, averaged over one hour. The blue line corresponds to the northern hemisphere and the red line to the southern hemisphere. The dotted lines indicate the ratio of the square of the mean electric field from Foster to the mean of the squared electric field from SuperDARN.

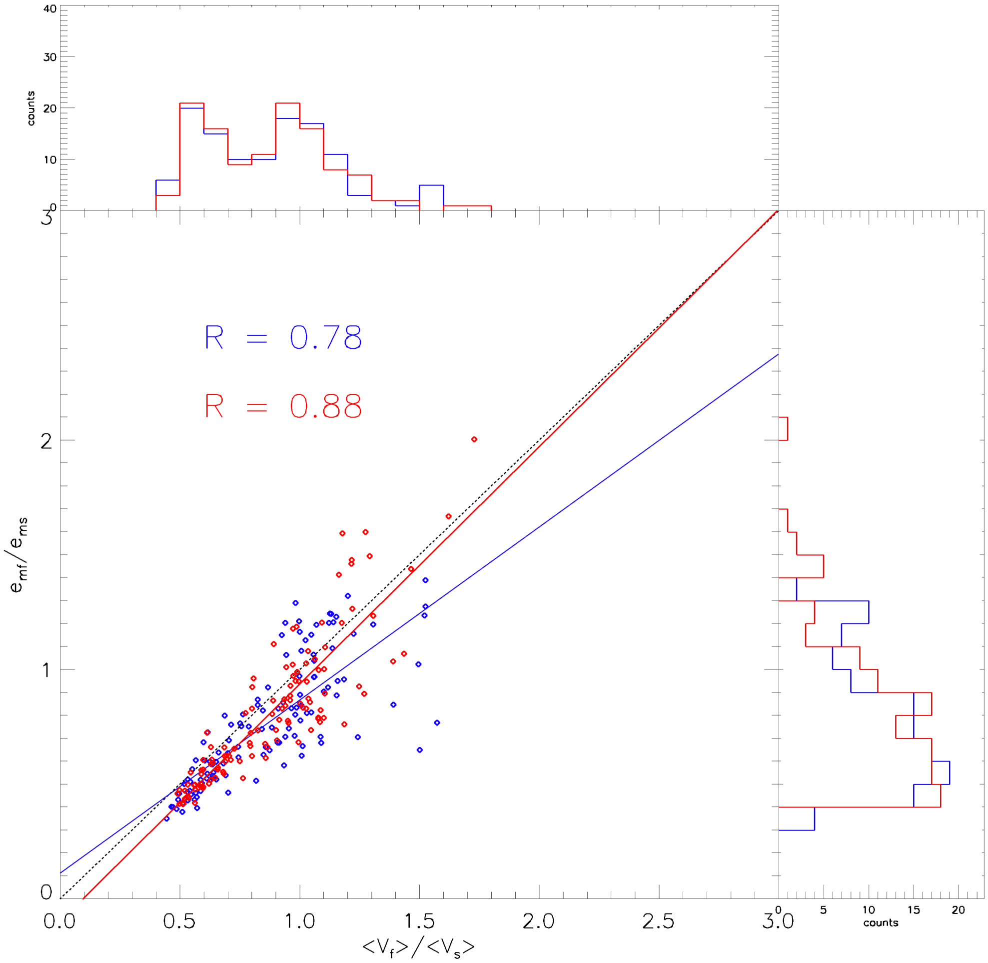


(b) The ratio of mean Joule heating from each model run plotted against the ratio of the square of the mean electric field from Foster to the mean of the squared electric field from SuperDARN.

Figure 6.8: A comparison of Joule heating in two model runs using SuperDARN data and using Foster's model driven by K_p. The histograms represent the number of data points in 0.2 wide bins, on both the x- and y-axes, to illustrate the distribution of values measured between the two models.



(a) The solid lines show mean neutral wind speed averaged hourly expressed as a ratio of the Foster and SuperDARN driven model runs. The blue corresponds to the northern polar region and the red to the southern polar region. The dotted lines indicate the correspond ratio of mean electric fields.



(b) A scatter plot comparing the ratios of mean Joule heating and mean electric field from the Foster and SuperDARN.

Figure 6.9: A comparison of neutral velocities from two model runs using Foster's electric field model and SuperDARN data. The format is the same as figure 6.8.

ratios e_{mf}/e_{ms} and $\overline{V_f}/\overline{V_s}$, with respective R values of 0.78 and 0.88 in the north and south. The plot also shows the ratios are very near to being equal, with gradients of 0.75 ± 0.06 (north) and 1.03 ± 0.05 (south) on the regression lines.

6.2.2 Discussion

The two simulations performed have shown the relationship $Q_J \propto e_s^2 + \sigma^2$ is a good representation of the heating in CMAT2. Figure 6.8(b) shows a strong correlation between the relative heating, $\overline{Q_f}/\overline{Q_s}$, and the electric fields $e_{mf}^2/(e_s^2 + \sigma_s^2)$ and the gradients are close to unity. This would indicate that the dependence of heating on the variability in the electric field is close to that proposed by Codrescu et al. (2000). The ratio of mean neutral wind speeds, $\overline{V_f}/\overline{V_s}$, between the Foster and SuperDARN models is found to be close to the ratio of mean electric field strengths, e_{mf}/e_{ms}, in figure 6.9(b). This shows that the neutral wind, whilst dependent on the mean electric field, is not strongly affected by its variability. Figure 6.9(b) shows this again, although there is not an exact match between the ratio of speeds and electric fields. This is most likely because the neutral wind is not driven directly by the electric field, but by ion-neutral drag. Additionally it is influenced by other forces, the most significant being the pressure gradient caused by radiative heating.

Although the results of these simulations are yet to be compared with any observational data, these comparisons between models serve to demonstrate the effects of different electric field inputs on the dynamics of CMAT2. These are necessary to address the effects of asymmetry between the northern and southern hemisphere electric fields, which are the subject of the following chapter.

Summary

Maps of electrostatic potential from SuperDARN have been used to drive CMAT2 during a five day interval at solar maximum during March equinoxes 2001. Section 6.1 studies the correlation between Foster's K_p-driven electric field model and the corresponding electric fields seen in SuperDARN. It is found that K_p is a poor proxy for the the ionospheric electric field. The K_p index is a measure of the maximum severity of geomagnetic activity within a three-hour window and so to use it as a proxy to derive an instantaneous magnetic field is inaccurate. Instead the IMF provides a better proxy for electric field behaviour at a much higher temporal resolution.

Section 6.2 addresses the issue of electric field variability and its effects when driving GCMs. The variable SuperDARN data is used to drive CMAT2 for a period of five days covering $16^{th} - 20^{th}$ March 2001, with the same electric field applied to both hemispheres. The model is run again for the same five day period but using Foster's electric field model for the corresponding K_p values, with otherwise identical input. The resulting Joule heating is found to follow the relationship, $Q_J \propto e_m^2 + \sigma^2$. The mean velocities, however, are found to be dependent only the mean electric field ($V \propto e_m$) and are not affected by variability. This suggests that for an empirical electric field model to produce realistic heating in GCM simulations it must both contain high resolution temporal variability, and be based on a proxy, or proxies, that accurately predict the strength of the electric field.

Chapter 7

Hemispheric Asymmetry in Ionosphere-Thermosphere Coupling

The previous chapter has explained the integration of SuperDARN-inferred electric fields into the CMAT2 GCM, and examined the merits of using these data over the model of Foster et al. (1986). Now the hemispheric asymmetries in the coupling of the ionosphere and thermosphere are examined. Firstly, the hemispheric deifferences in observations the northern- and southern-SuperDARN instruments are considered. This is followed by a modelling study that uses these data to study their energy and momentum input to the neutral atmosphere, which is then compared to ground- and space-based observations.

The weakness of SuperDARN data is that the periods of low back-scatter make it reliant on an empirical, IMF dependent climatological model. For a modelling study, however, it is possible to select periods of data where there is reasonably high back-scatter in both hemispheres for a period of several days. The SuperDARN data can then be assumed to be the best available representation of the ionospheric electric field under these circumstances. The magnetic field used by CMAT2 is a simplified version of the IGRF, which is a spherical harmonic expansion of the Gauss coefficients up to only third order. The reason for this, as explained in section 4.5, is that if the full IGRF is used in CMAT2, it causes the flux tubes used by the Global Ionosphere-Plasmasphere model to converge at certain positions and greatly reduces the resolution of this module. This does still, however, provide a magnetic field which is approximately a tilted dipole and offset from the geomagnetic poles, although the offset does not match the true IGRF. The positions of the magnetic dip poles in CMAT2 and the value from the 11^{th} generation are shown in table 7.1

	North dip pole		**South dip pole**	
	Latitude	**Longitude**	**Latitude**	**Longitude**
18^{th} order	80.97°	-109.64°	-64.66°	138.30°
3^{rd} order	80.52°	153°	-60.08°	144°

Table 7.1: Geographic coordinates of the north and south dip poles from the 11^{th} generation IGRF to 18^{th} and 3^{rd} order, for the year 2010.

7.1 Hemispheric Dependence of SuperDARN on IMF Conditions

Before addressing the asymmetry between ionosphere and thermosphere it is worth looking first at the asymmetry within the data provided by SuperDARN in each hemisphere. The arguments made by Codrescu et al. (2000), which were addressed in the previous chapter, are based on the assumption that the electric field variability followed a normal distribution. With the available SuperDARN data it is possible to test this assumption somewhat by considering the cross cap potential difference in each data set provided. It is known that the high latitude electric field correlates with the strength and orientation of the IMF vector and so the values of $\Delta\Phi_{pc}$ for selected electrostatic potential maps from SuperDARN have been binned accordingly. The SuperDARN data for the years solar maximum years 2000-03 are selected and divided in to 24 bins according to the eight IMF sectors and three levels of magnitude: $|\mathbf{B}| < 5.0$nT, 5.0nT$\leq |\mathbf{B}| < 7.5$nT and 7.5nT$\leq |\mathbf{B}|$. Data sets are selected in the NH where there are $\geq$200 data points used in the fit. Due to lack of measurements in the SH, this constraint is relaxed and only data sets containing no data are excluded. Although not ideal, restricting the SH data to ≥ 200 data points does not result in enough measurements in each IMF bin to produce a meaningful distribution. Because the OMNI 1-minute data set is used to determine the IMF sector, data are only included where the IMF sector has remained constant for at least ten minutes. This condition is used to avoid the discrepancy between the OMNI estimate and the observed SuperDARN field, which acts to smooth out the observed differences between potential maps in different IMF sectors. $\Delta\Phi_{pc}$ represents a large scale picture of each electrostatic potential map and does not contain a measure of small grid-scale variability and also is not a direct measure of the electric field. However, in order to bin the data as described here it is necessary to quantify each electrostatic potential map by a single value. The electric field is the gradient of the potential and so the magnitude of $\Delta\Phi_{pc}$ are closely related. It is therefore assumed that $\Delta\Phi_{pc}$ will give a good representation of the type of distribution that classifies electric field variability.

For each bin the cross-cap potential is determined and a histogram created for each hemisphere. These distributions are shown for the three levels of IMF strength in figures 7.1, 7.2 and 7.3. These three levels are selected because they produce approximately similar numbers of counts in each histogram, which are displayed in table 7.2 for each hemisphere.

	Northern hemisphere			**Southern hemisphere**		
sector	**low**	**medium**	**high**	**low**	**medium**	**high**
0	5299	3695	4070	7171	4932	5327
1	7533	6813	5528	10374	8875	7242
2	13307	11956	8340	17624	15673	10805
3	8055	6618	5305	10920	8444	8121
4	5208	3969	4125	6940	5164	5816
5	7859	6203	4977	10811	8136	6710
6	11740	8014	6400	15700	12111	9080
7	7772	4956	5326	10365	6984	7297

Table 7.2: Number of data-sets in each bin that were used to produce the potentials in figures 7.1, 7.2 and 7.3. The *low*, *medium* and *high* headings correspond to the three levels of IMF magnitude, separated by 5.0 and 7.5nT.

Above each of the histograms in figures 7.1, 7.2 and 7.3, are printed the mean and mode of each distribution. It is clear from inspection that none of the histograms are well represented by a normal distribution. In fact, depending on the IMF bin, the types of distribution vary quite significantly. In almost all cases, and particularly under weaker conditions in figures 7.1 and 7.2, the distributions show a very sharp peak. Another feature seen in some of the distributions, particularly for strong IMF (figure 7.3), are histograms where there are more counts to one side of the peak. These more closely resemble a log-normal distribution. In some cases both of these features are present, for example under

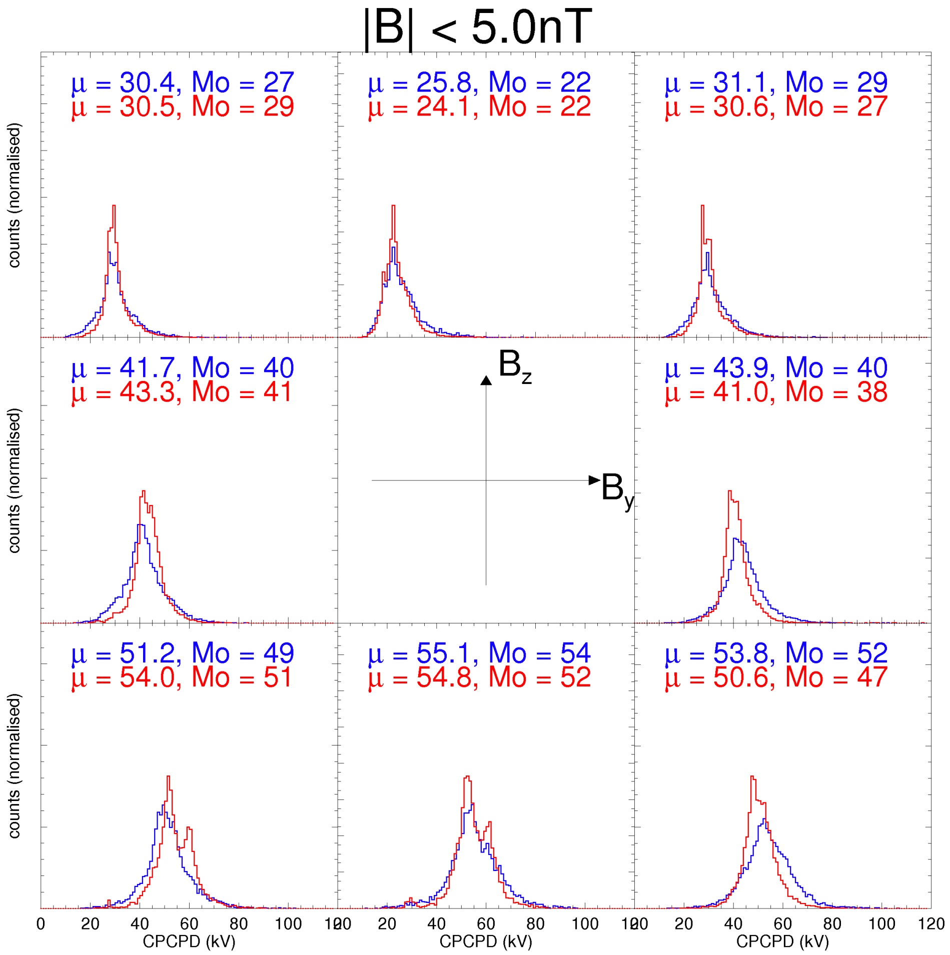


Figure 7.1: Histograms showing the number of counts for a given cross-cap potential recorded by SuperDARN as a function of IMF sector for $|\mathbf{B}| < 5.0$nT. The NH is shown in blue and the SH in red. The SH histograms are normalised such that their peaks are equal. The NH histograms are normalised such that their area is equal to the corresponding SH histogram. The mean and mode of each distribution are printed above.

strong IMF conditions and northward B_z in figure 7.3. An interesting study would be to fit the appropriate probability density function to each of these distributions and derive the effect of variability on Joule heating, in the same way that Codrescu et al. (2000) didi using a normal distribtuion. However, due to the complicated and diversity of these distributions this was not acheived. In some cases the distributions appear bimodal with two distinct peaks. This is particularly apparent for $B_y < 0$ and $B_z < 0$ (bottom left, sector 5) in figure 7.2 and appears in several SH distributions. If there are many back-scatter measurements

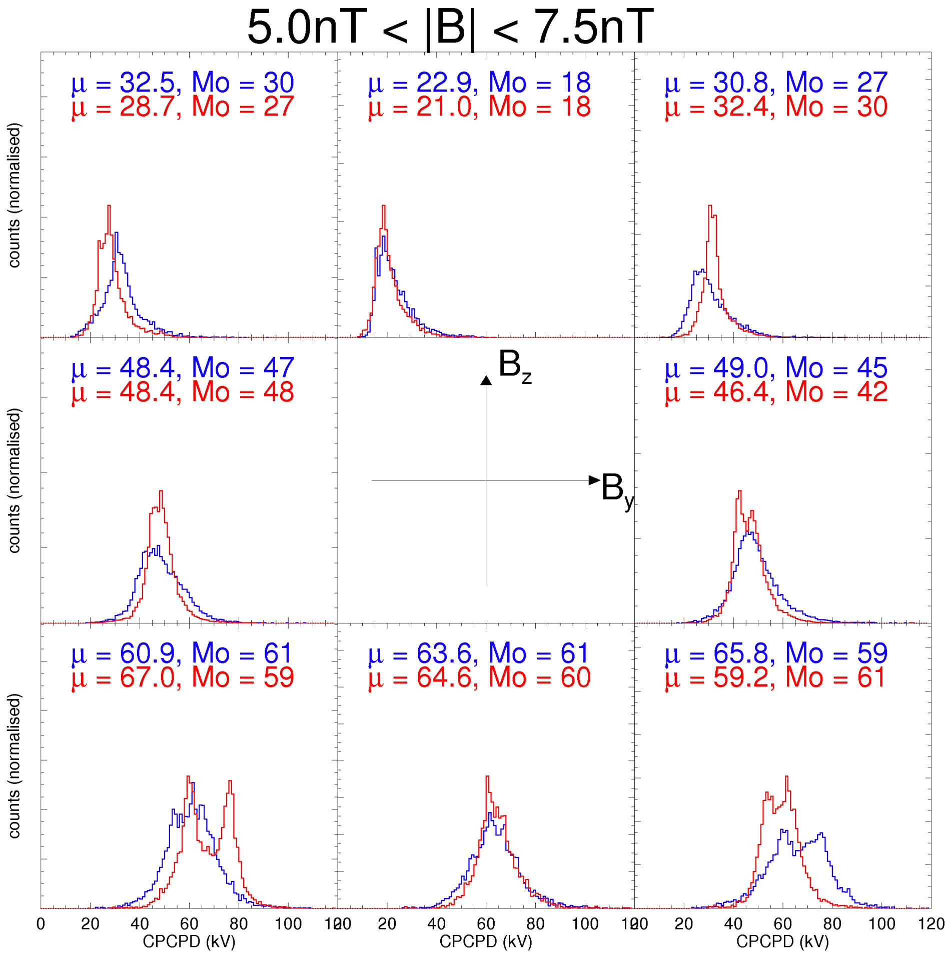


Figure 7.2: The same format as figure 7.1 for $5.0 \leq |\mathbf{B}| < 7.5$nT.

used to fit a convection pattern the value of $\Delta\Phi_{pc}$, in a given bin, would be expected to fall close to the mean value for that bin, however in the absence of measurements it would fall near the value determined from the empirical model. If these two values differ then one would expect to see two peaks when the low- and high-back-scatter data sets are binned together, as they have been here in the SH. Because the more strict condition that over 200 data points must exist is applied in the NH, the same effect is not seen in that hemisphere.

In general these distributions show the expected B_z dependence on the cross-cap potential; when it is negative $\Delta\Phi_{pc}$ increases, and this effect is greater as the magnitude of $\mathbf{B}$ increases. For the three increasing levels of IMF magnitude $\Delta\Phi_{pc}$ is greater by approximately $30, 40$ and 50kV between B_z negative and positive. There is also a B_y dependence,

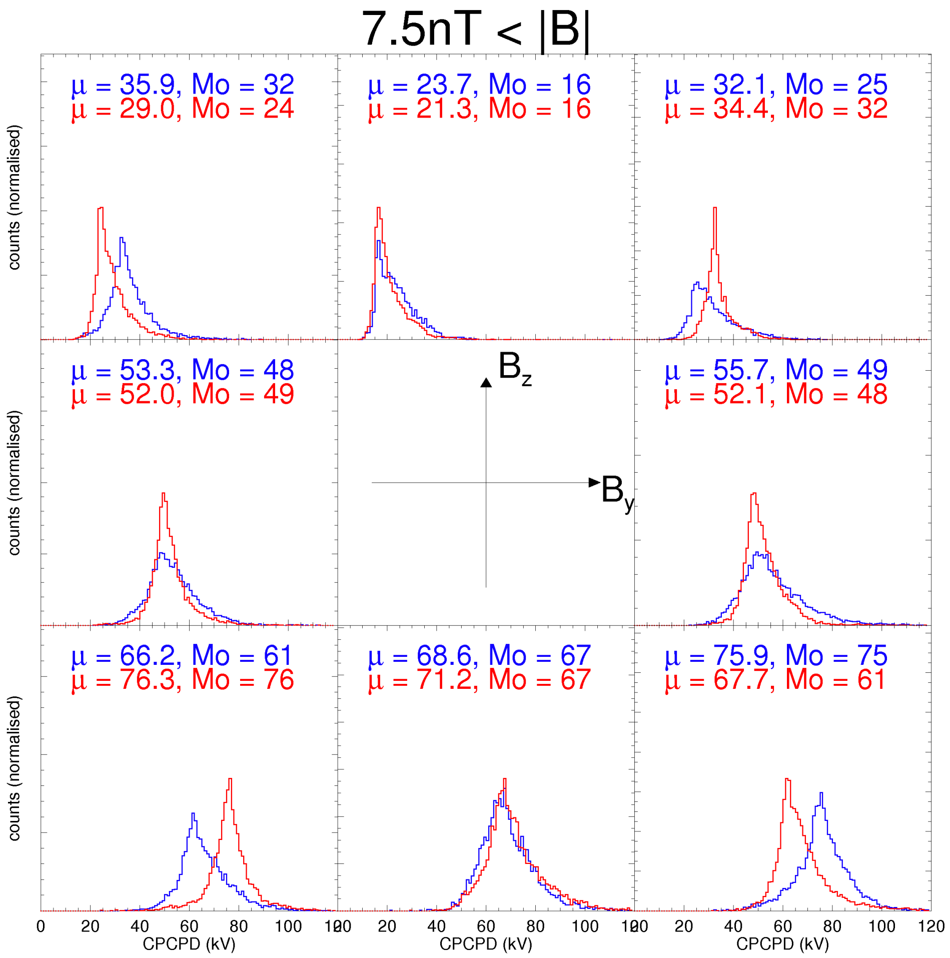


Figure 7.3: The same format as figure 7.1 for 7.5nT$\leq$ $|\mathbf{B}|$.

which becomes more apparent when the IMF magnitude is strongest (figure 7.3). For $B_z < 0$, B_y positive (negative) produces a stronger electric field in the NH (SH), as described at the beginning of this section. For $B_z > 0$, this effect appears to be reversed in each hemisphere. The histograms also suggest that the distributions are not normal, but instead peaked and heavy-tailed. They generally show a heavy right-tail (more counts above peak value), an effect which is clearer for large $|B|$, in figure 7.3.

In the case of the most sharply peaked distributions, for example the top three sectors in figure 7.1, the values generally fall close to the mean. The deviation from the mean is less than that of a normal distribution and, by Codrescu's argument, we would expect this variability to have less of an effect for a peaked distribution than for a normal distribution.

For the very heavy right-tailed distributions (figure 7.3), and particularly the bimodal distributions, we would expect the opposite to be true because greater values, with respect to the mean, will be more commonly observed than from a normal distribution.

It is clear that there is a hemispheric dependence on the strength of the averaged electric fields in each hemisphere, which becomes more significant for strong IMF and when B_z is negative. Additionally, the variability appears to show some dependence on IMF strength; for the strongest IMF conditions (figure 7.3) the distributions appear more skewed, which suggests the difference between $\overline{e^2}$ and $\overline{e}^2$ will be larger than predicted by Codrescu et al. (2000). For weaker conditions it is heavy-tailed suggesting the difference between $\overline{e^2}$ and $\overline{e}^2$ will be less. The results in the previous chapter, for general solar wind conditions, showed Codrescu's argument to be well supported. This may be the result of the two effects cancelling each other out.

7.2 Electrostatic Potential Maps

Further to the histograms in the previous section, it is possible to learn more about the hemispheric B_y dependence of SuperDARN from the electrostatic potential distributions. The SuperDARN data are selected using the same criteria as in the previous discussion; ≥ 200 data points in the NH, > 0 in the SH and stable IMF for at least 10 minutes from the years 2000-03. These are then binned in the same way according to IMF orientation and magnitude. The electrostatic potential distribution is plotted over a grid of 2° magnetic latitude and 18° magnetic longitude using the binned data sets. In each bin, the mean potential and standard deviation in each grid cell are calculated, which are plotted in figures 7.4, 7.6 and 7.8 (NH) and 7.5, 7.7 and 7.9 (SH).

Figures 7.4 to 7.9 show the mean and standard deviation from the SuperDARN potentials binned by IMF sector and strength. With an increase in the magnitude $|\mathbf{B}|$, and with more strongly southward B_z, an increase in $\Delta\Phi_{pc}$ is observed. This is consistent both with established behaviour and with the results in the previous section. A B_y dependence is also observed, where a positive B_y in the NH favours a stronger $\Delta\Phi_{pc}$ and a negative B_y has the same effect in the SH. Additionally, a dependence of the shape of the distribution on B_y and B_z can be seen. Multiple convection cells during B_z positive can be seen throughout both hemispheres at all IMF strengths. These are most clear in figures 7.5, 7.6 and 7.8 due to the level of the contour chosen in the plot. There is a general increase in size of the

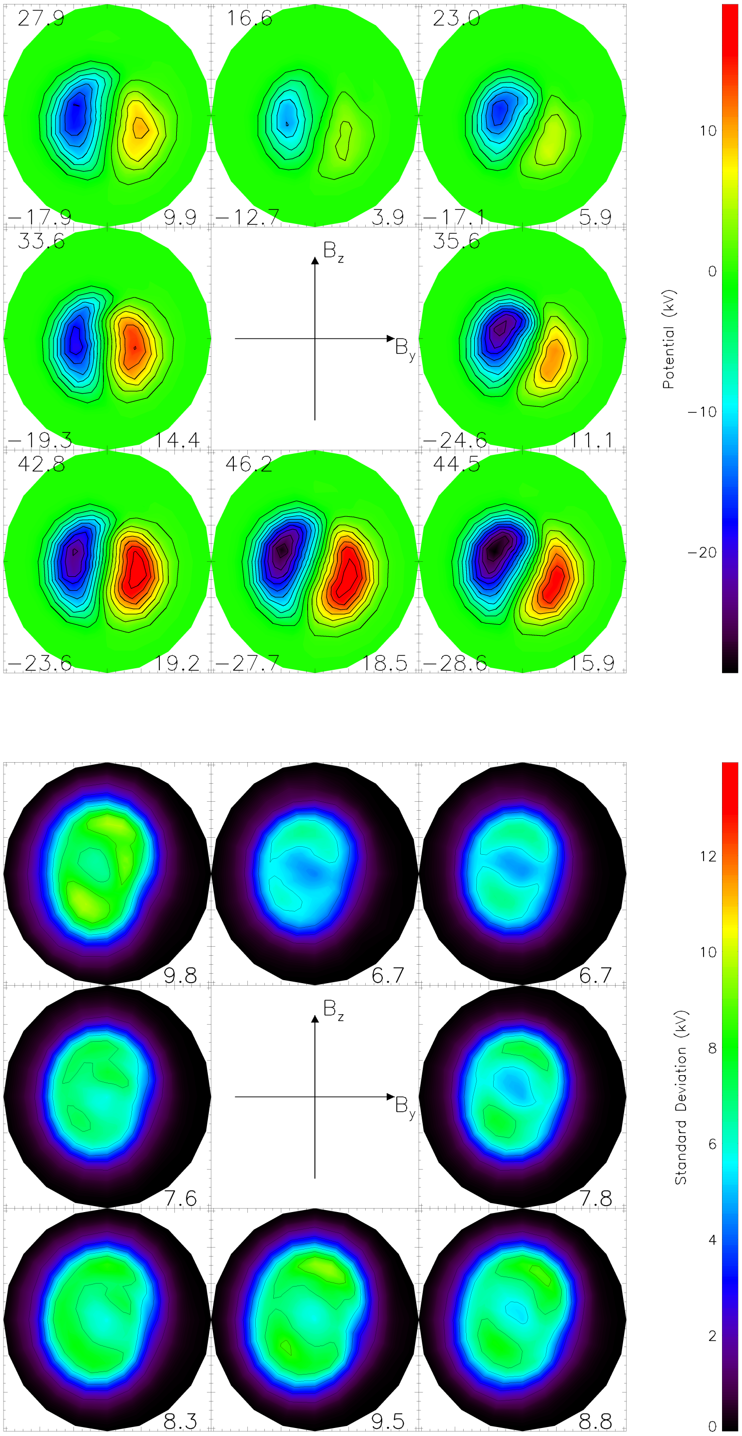


Figure 7.4: Mean and standard deviation of electrostatic potential in the NH over a grid of magnetic latitude and longitude, with 12MLT to the top of the plot. The data are binned according to IMF sector, for $|\mathbf{B}| \leq 5.0$nT using the same selection criteria as the previous section. The printed numbers represent the cross cap potential and maximum and minimum values in the top plot and the maximum value in the bottom plot.

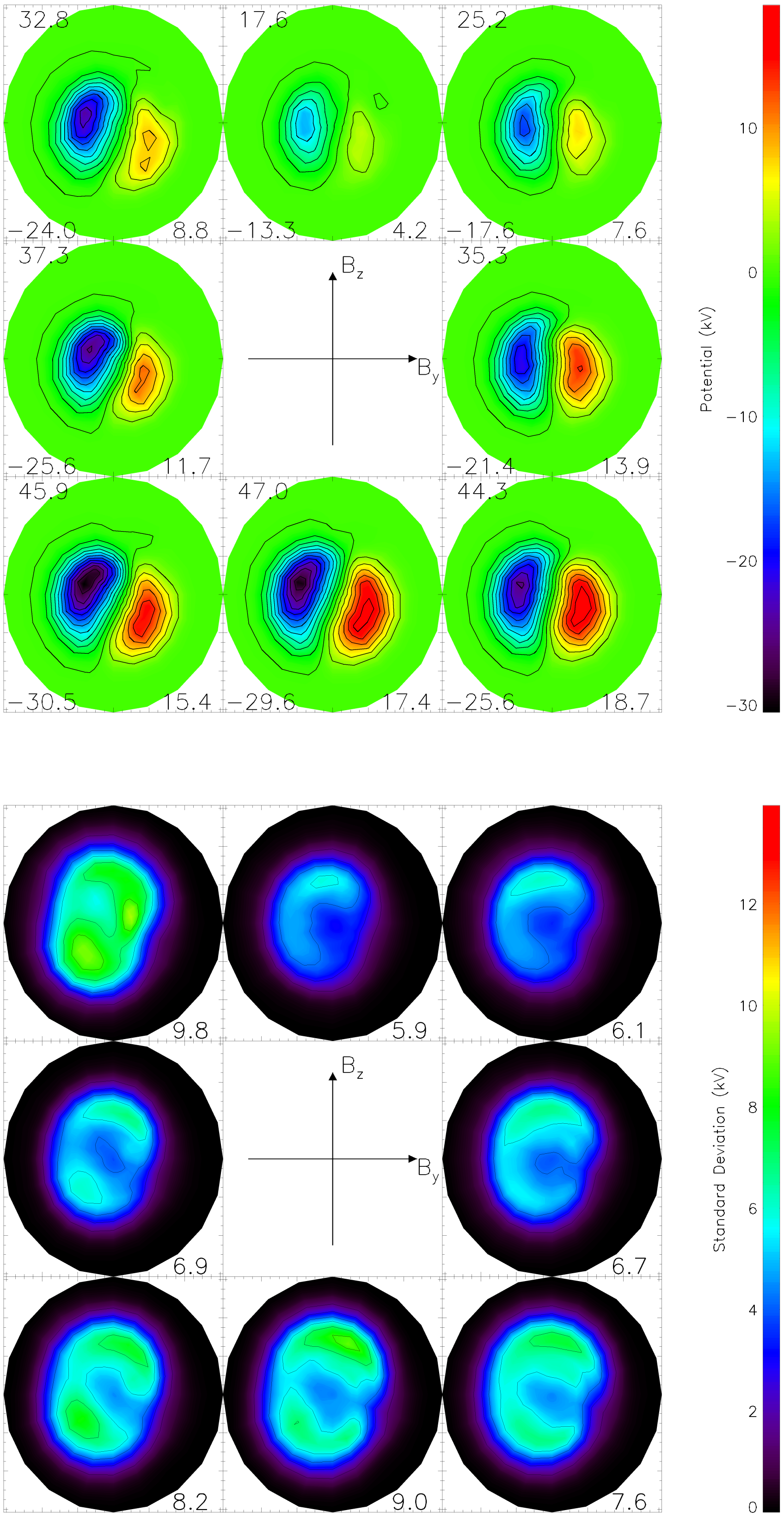


Figure 7.5: The same format as figure 7.4, for the SH.

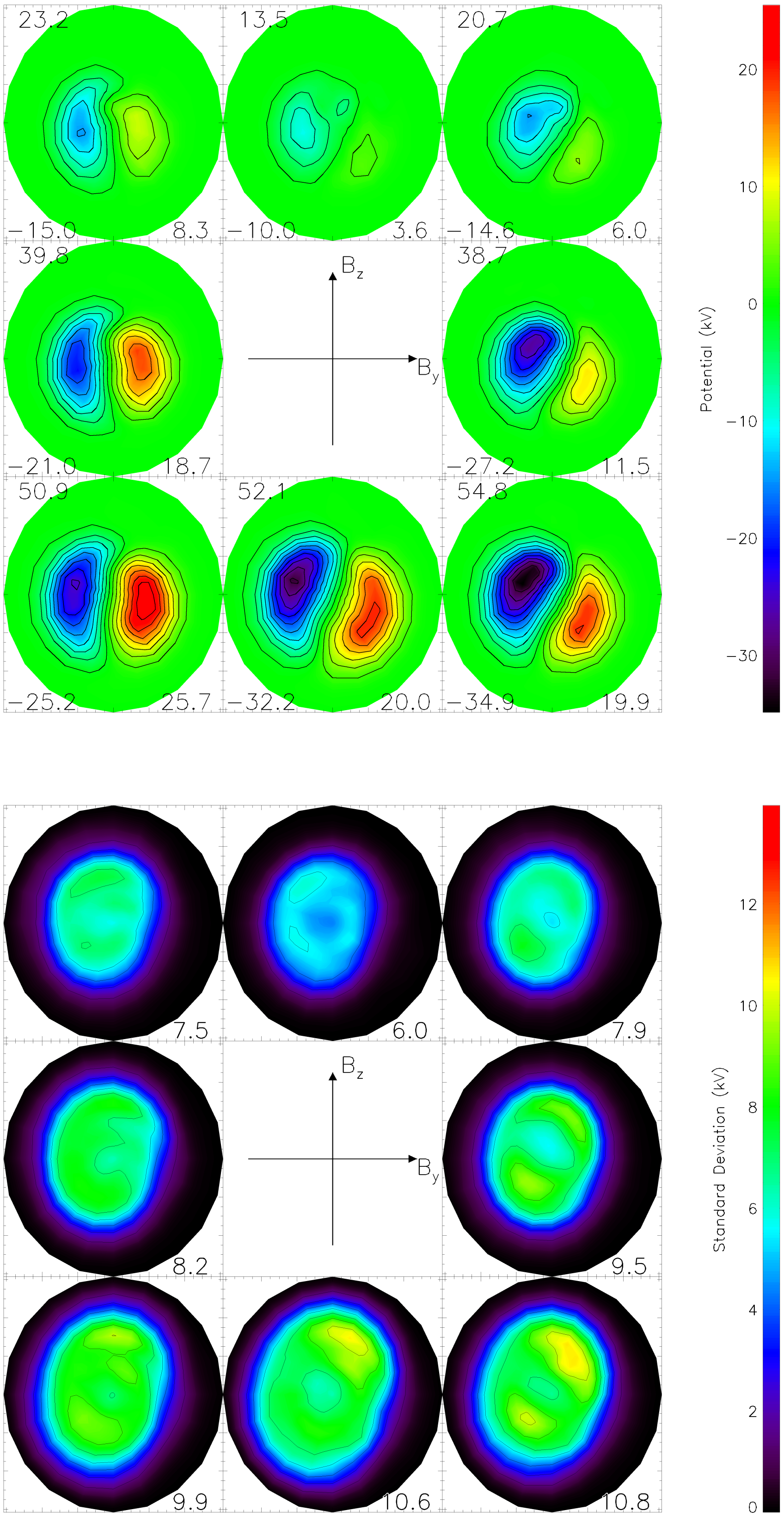


Figure 7.6: The same format as figure 7.4, for the NH when 5.0nT$< |\mathbf{B}| \leq$ 7.5nT.

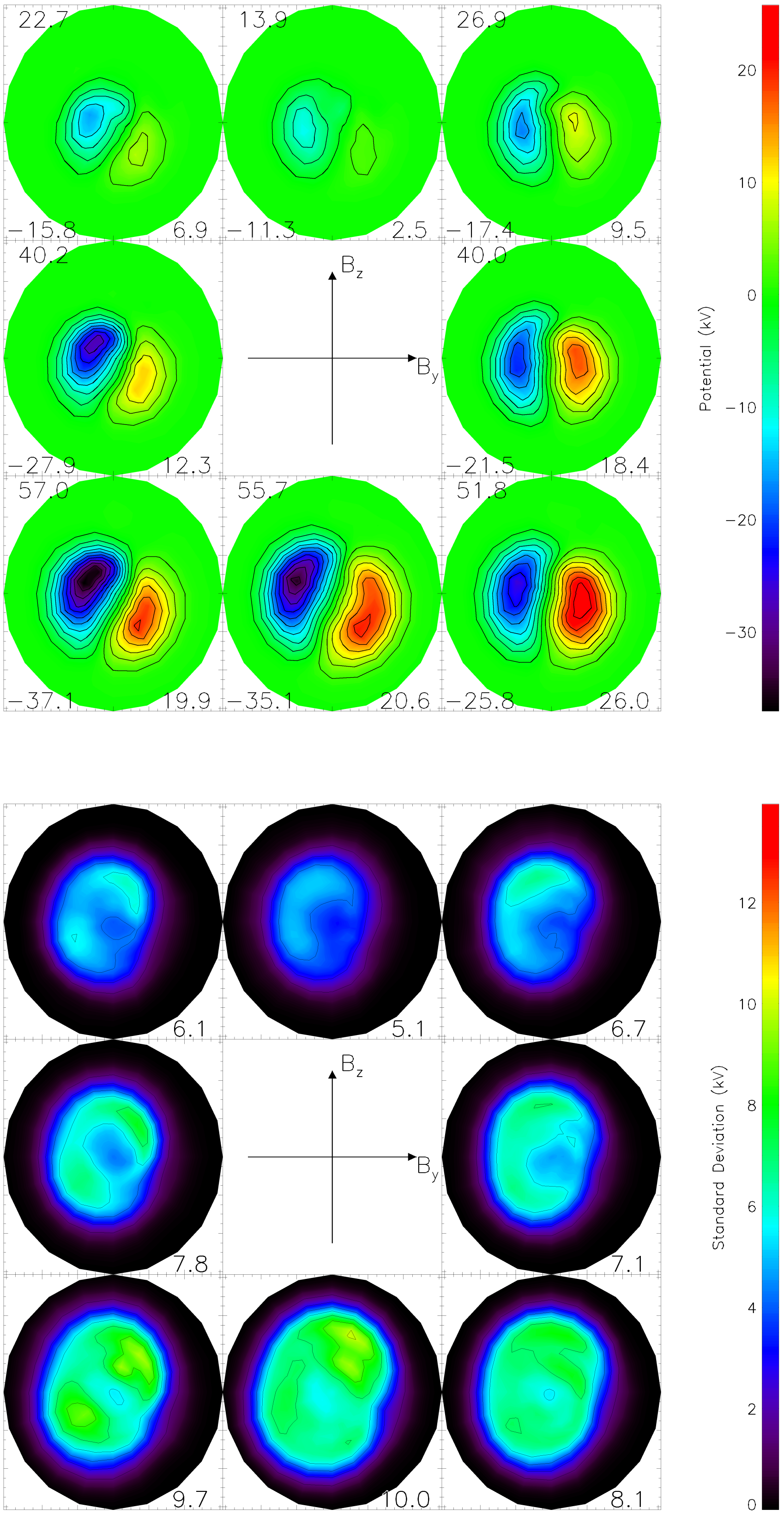


Figure 7.7: The same format as figure 7.4, for the SH when 5.0nT$<|\mathbf{B}|\leq$ 7.5nT.

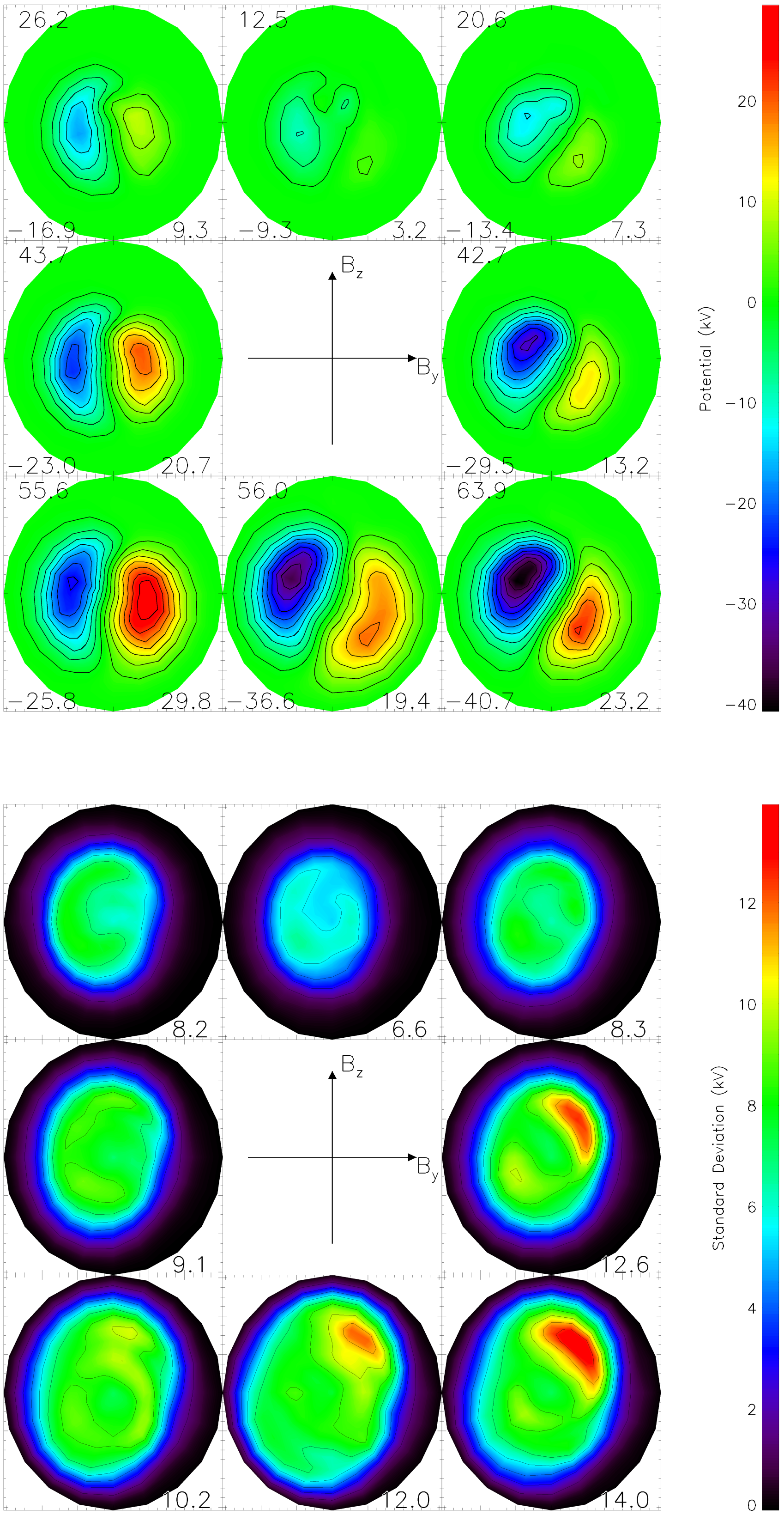


Figure 7.8: The same format as figure 7.4, for the NH when 7.5nT$\leq |\mathbf{B}|$.

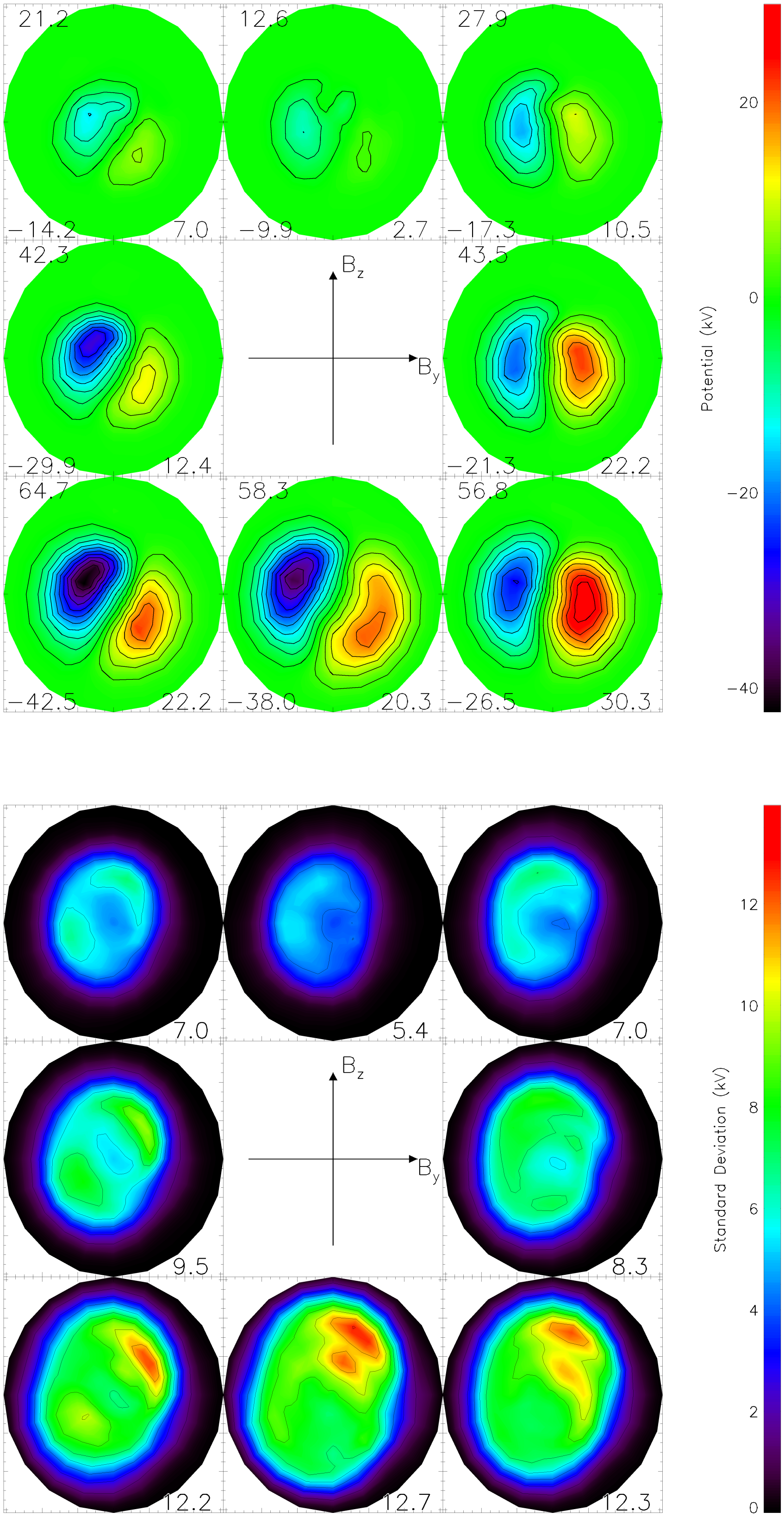


Figure 7.9: The same format as figure 7.4, for the SH when 7.5nT$\leq |\mathbf{B}|$.

overall convection pattern as B_z becomes more strongly negative, as was observed in the expansion of the HMB in figure 6.5. A further effect of B_y on the potential distributions is observed in that it acts to "twist" the convection cells. For example, when $B_y > 0$ in the NH, the cells are rotated clockwise, relative to their position when $B_y < 0$. This effect is also present for all IMF strengths. In the SH, the effect is seen but with the opposite dependence on the sign of B_y. The maps of standard deviations in potential are shown below in each of the figures 7.4 to 7.9, which follow the general trend of the mean potential values; there is an increase in variability with negative B_z and large $|\mathbf{B}|$, whilst B_y positive (negative) corresponds to an increase in variability in the NH (SH). The variability generally shows a peak between approximately 06 and 12MLT for the most southward IMF conditions, otherwise a peak is less well defined. Under southward IMF in the SH, and particularly for strong IMF values (figure 7.7 and more so in figure 7.9) there is a double peak, again between 06 and 12MLT.

7.3 CMAT2 Simulations

An advantage of using SuperDARN data to drive CMAT2, over the Foster and Weimer models, is that it provides separate coverage of each of the northern and southern polar caps. GCMs with MHD driven electric field input do also possess this asymmetry and have been applied to study various aspects of the hemispheric differences in convection. One such model is the Coupled Magnetosphere-Ionosphere-Thermosphere (Wiltberger et al., 2004), which consists of the Thermosphere-Ionosphere-Electrodynamics General Circulation Model (TIE-GCM) with the LFM magnetosphere coupled to the upper boundary. Förster and Cnossen (2013) used two versions of this model, one with a simple dipole magnetic field and a second with the IGRF, the former being effectively symmetric in each hemisphere and the latter being asymmetric. They studied the effects of both the differing flux densities and the different offset between the magnetic and geographic poles and concluded that, in the case of the IGRF, the model results showed significant hemispheric differences in neutral wind circulation at high magnetic latitudes. Conversely, the dipole model resulted in generally symmetric behaviour. In this chapter, an alternative approach is applied to the same problem of asymmetry in ionosphere-thermosphere coupling. Model simulations are performed using different electric field input based on SuperDARN observations; an asymmetric model run simply uses the SuperDARN data in each hemisphere to drive the model, whilst a symmetric run uses the same (northern hemisphere) data to drive both hemispheres.

In order to study the effects on the neutral atmosphere that result from the different electric field patterns in each hemisphere, two model runs have been performed using CMAT2. In the first, SuperDARN observations of the NH are used to drive the model in *both* the NH and SH, similar to the way the Foster model was used, whilst the second model run uses independent measurements for each hemisphere. The model is run for two five day periods, where SuperDARN was receiving a large amount of back-scatter in both hemispheres. These periods cover the 16-20 March 2001 and 14-18 September 2002 during solar maximum, which are close to opposite equinoxes, so that any seasonal differences should be minimised. The IMF parameters for these periods, as measured by the ACE and WIND spacecraft, are shown in figure 6.6 and 7.10. The purpose of using these hemispherically-symmetric and -asymmetric electric field inputs to CMAT2 is to observe the resulting symmetries/asymmetries in the neutral atmosphere and to establish how the two are linked.

7.3.1 Magnitude and Extent of the Electric Field

Figure 7.11 contains SuperDARN electrostatic potential data from each of the five-day simulation periods; March 2001 (left) and September 2002 (right). Each plot shows the maximum and minimum hemispheric potentials, from each two-minute time-step, binned according to the IMF sector from the OMNI data-set. The observed values in both cases exhibit the electric field B_y dependence described at the beginning of this section, where the NH dawn cell (maximum potential) is enhanced when $B_y < 0$ and the dusk cell (minimum potential) enhanced when $B_y > 0$. The opposite is true in the SH, where the B_y dependence is reversed. It is also apparent that, in both five day simulations, the absolute potential in the dusk cell (the negative values) is typically greater than that of the dawn cell. This is observed in both hemispheres. Whilst the maximum and minimum values in each plot show a dependence on both B_y and B_z, the difference between the maximum and minimum values ($\Delta\Phi_{pc}$) in each hemisphere is far more dependent on B_z alone. This effect can be seen in figure 6.4, where $\Delta\Phi_{pc}$ in Foster and Weimer's models, as well as both hemispheres of SuperDARN, show no significant B_y dependence.

Figure 7.12 shows the difference between the average electric field magnitude in each hemisphere in the region $|\phi_m| > 80°$ over both five day simulations. In each plot there

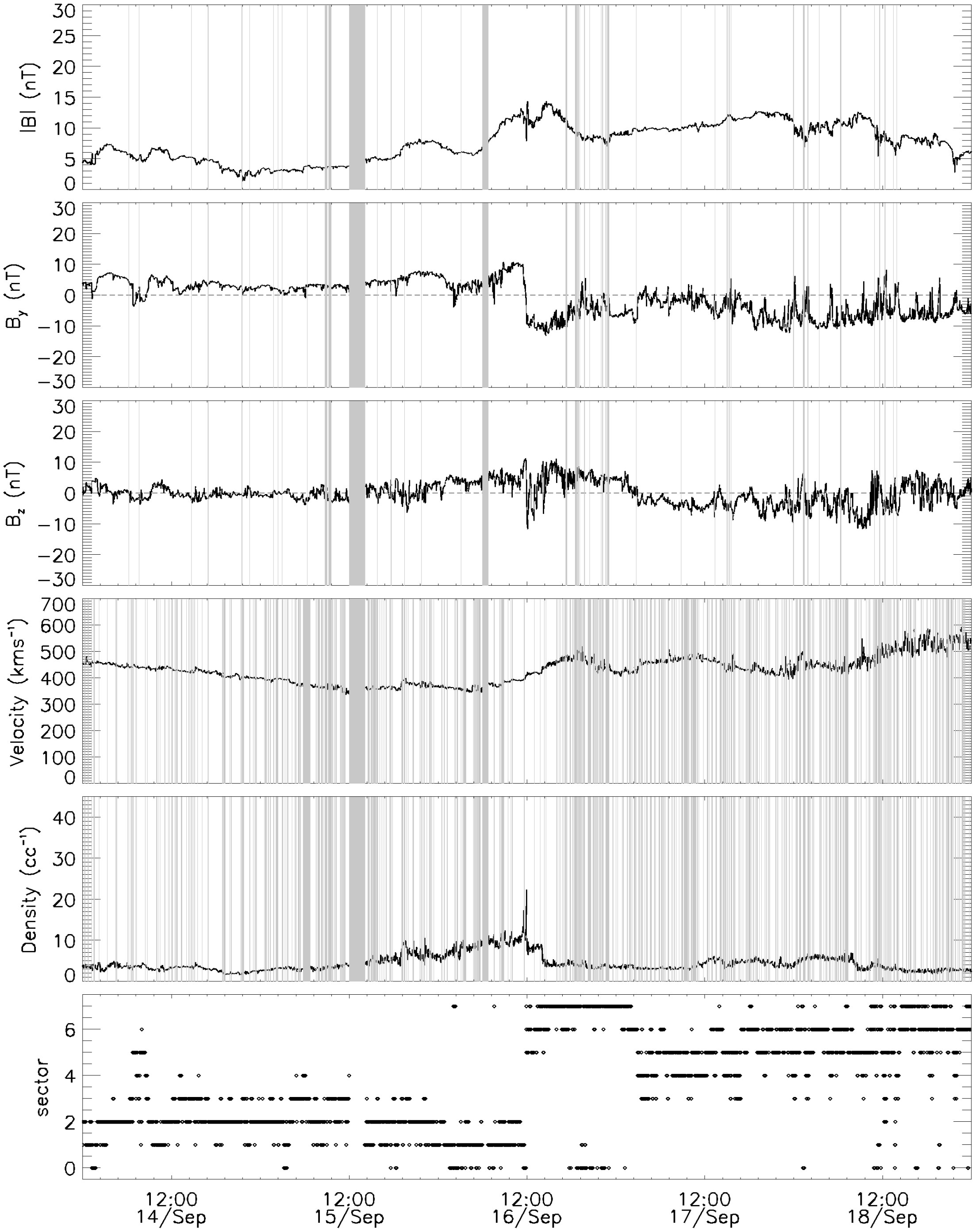
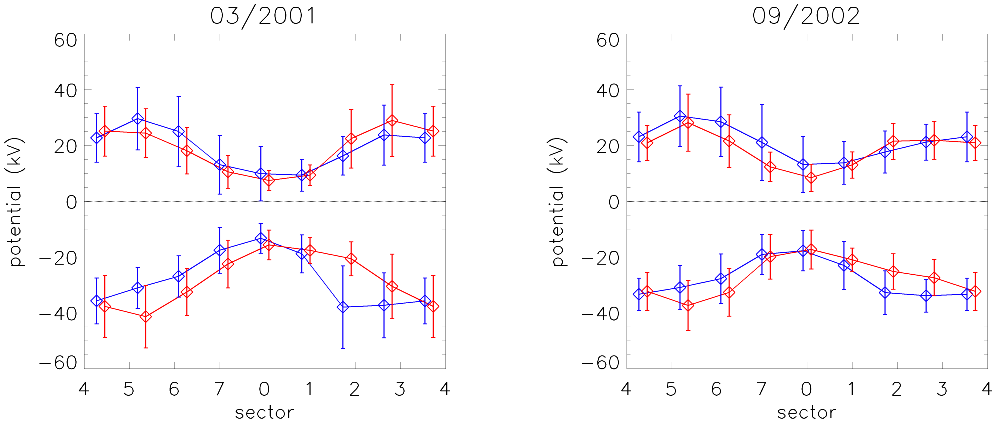


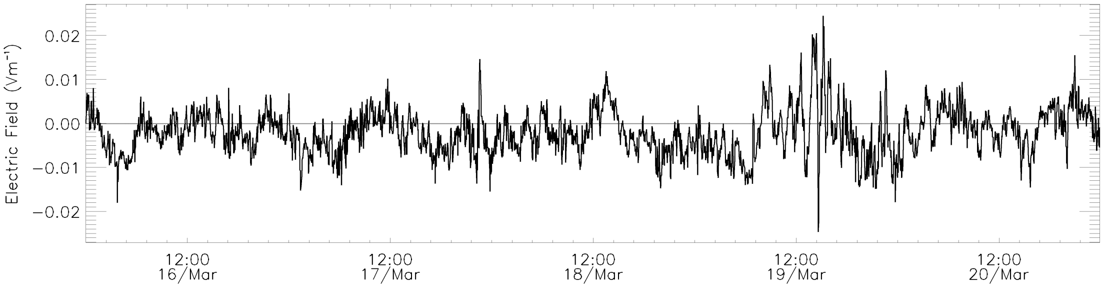
Figure 7.10: Solar wind and IMF parameters for 14 − 18 September 2002, as measured by the ACE and WIND spacecraft. The Same format as figure 6.6.



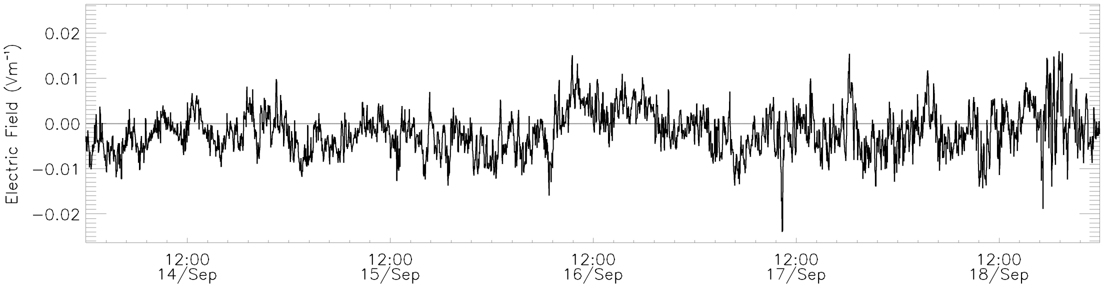
(a) March 2001 simulation.

(b) September 2002 simulation.

Figure 7.11: Maximum and Minimum electrostatic potential in the NH (blue) and SH(red), as a function of IMF sector for each five day simulation.



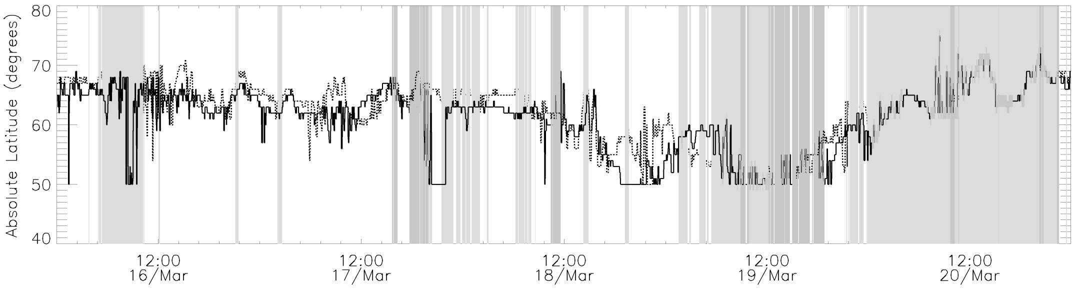
(a) 16-20 March 2001.



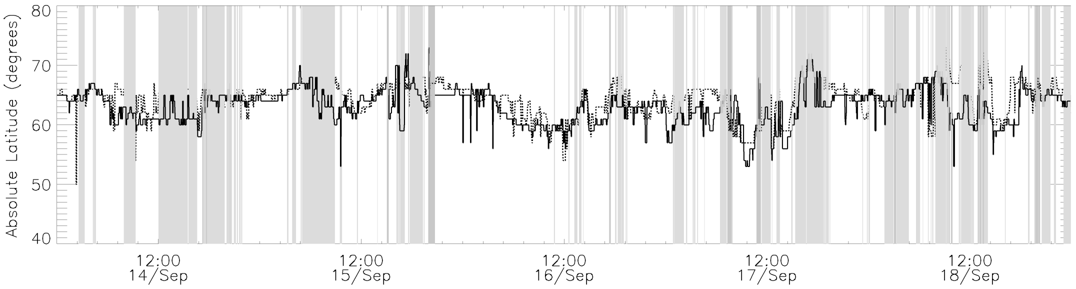
(b) 14-18 September 2002.

Figure 7.12: Difference in average electric field magnitude in the region $|\phi_m| > 80°$ between the NH and SH over both five day simulation intervals.

appears to be little difference between the two hemispheres, although the negative values dominate slightly suggesting a stronger electric field in the SH. This is in agreement with the results in figures 7.4(a) to 7.9(a), which show the average cross cap potential is marginally greater in the SH. The plots in figure 7.12 appear not to correspond to any changes in IMF conditions seen in figures 6.6 and 7.10.



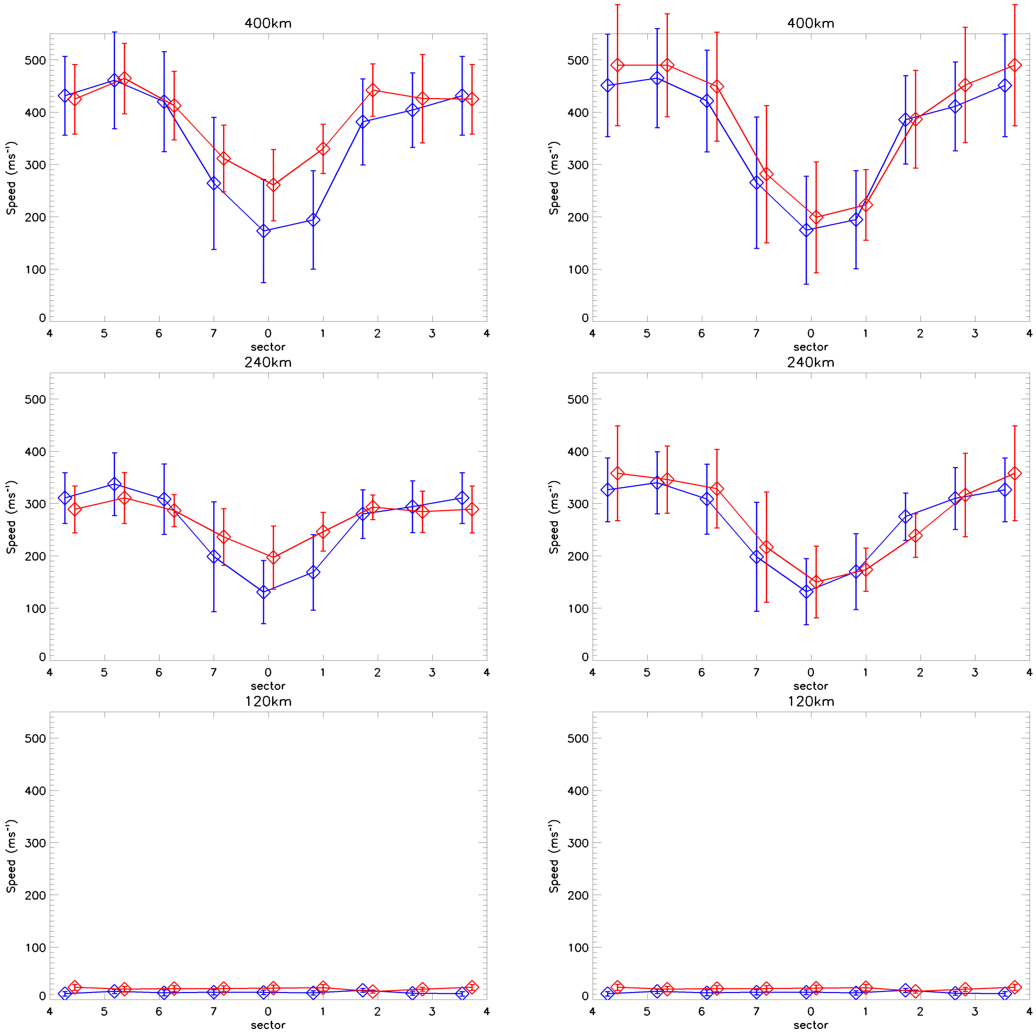
(a) 16-20 March 2001.



(b) 14-18 September 2002.

Figure 7.13: The position of the HMB over both five day simulation intervals. The solid line represents the NH and the dotted line the SH. The greyed-out areas indicate times where the number of data points in either hemisphere was fewer than 200.

Figure 7.13 shows the latitudinal extent of the HMB in each hemisphere over both the simulation intervals. In the March 2001 simulation, in both hemispheres the HMB appears to remain at $|\phi_m| \sim 65°$ during the quiet period over the first ∼60 hours. During the period of high geomagnetic activity in the final 60 hours, both show a significant expansion of the size of the convection pattern. The HMB in both hemispheres extends as low as 50° for a prolonged period, although this corresponds to a time when there is generally a lack of data points produced from the SuperDARN observations. In the September simulation, there are generally fewer periods where the number of data points is less than 200 and the HMB, in both hemispheres, varies little between $60 - 70°$. Overall, both simulations show little difference between the NH and SH location of the HMB, as a function of time.



(a) Model run performed using asymmetric electric fields.

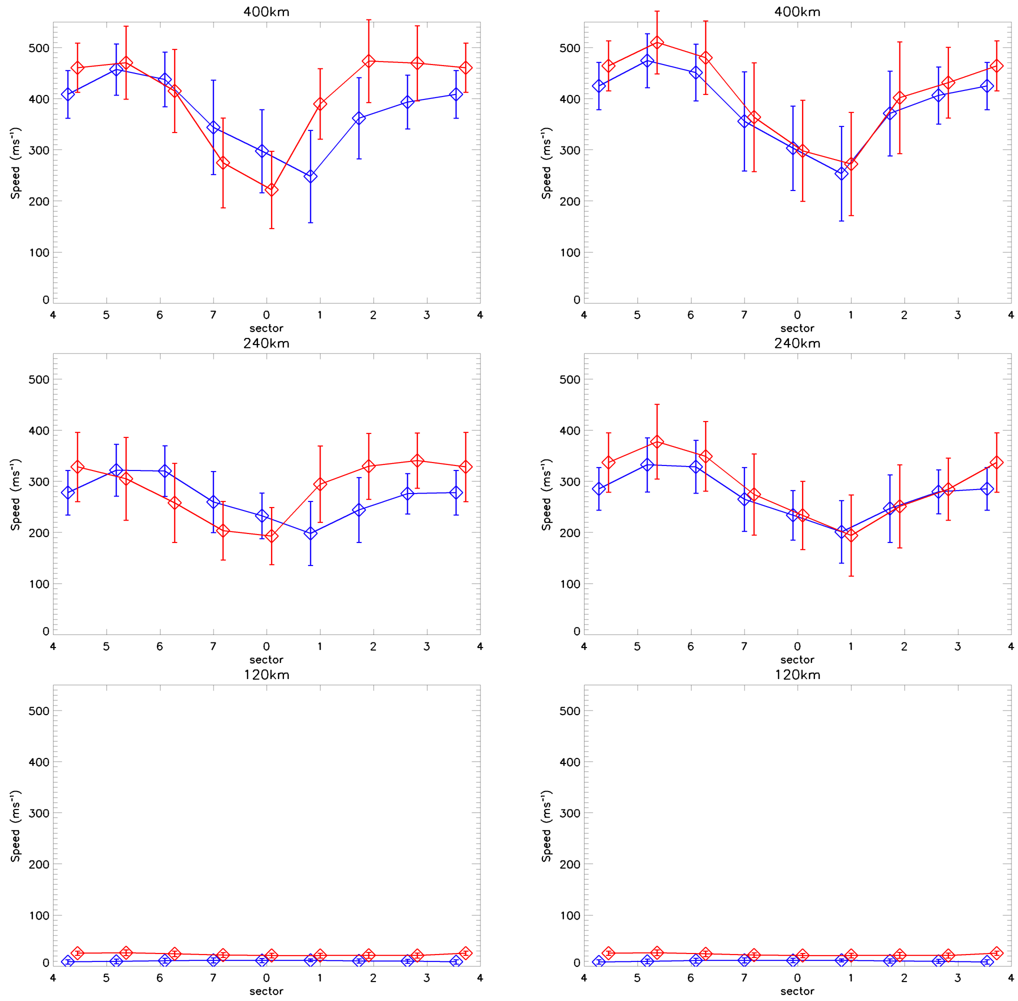
(b) Model run performed using identical electric fields in both hemispheres.

Figure 7.14: Neutral wind speeds average over the northern and southern polar regions, $|\phi_m| > 80°$, from the 16-20 March simulation. and binned according to IMF sector. Again, blue and red represent the NH and SH, respectively, and the dotted lines are the standard deviation in each bin.

7.4 CMAT2 Results

7.4.1 Neutral Wind Speeds

For each of the simulations, using both the symmetric and asymmetric electric fields, the neutral velocities in the region $|\phi_m| > 80°$ in each hemisphere are binned according to IMF sector and averaged. Figure 7.14 shows the results at 120km, 240km and 400km. In both the asymmetric (left) and symmetric (right) simulations, the model displays a clear correlation between B_z and neutral wind speeds, where the average wind magnitude in



(a) Model run performed using asymmetric electric fields.

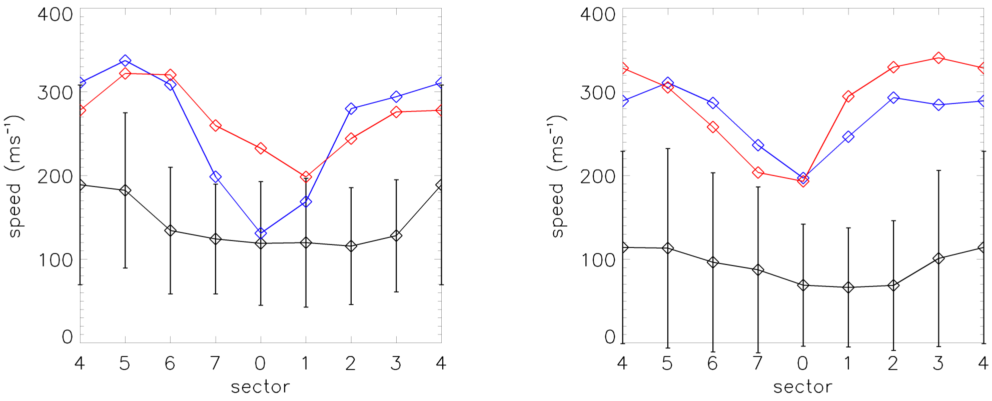
(b) Model run performed using identical electric fields in both hemispheres.

Figure 7.15: The same format as figure 7.14, for the 14-18 September 2002 simulation.

sector 4 is approximately double the value in sector 0. This effect is seen most clearly at 240km and 400km, below which the velocity is very low. At 120km the neutral mass is too great to be affected by ion convection, resulting in no IMF dependence at this altitude. In both models the SH speeds are typically very slightly greater, which is likely to result from both the greater value of $|\mathbf{B}|$ (from the IGRF) and of $|\mathbf{E}|$ (from SuperDARN) in this hemisphere. There is a small difference in how the models respond to different signs of B_y at the higher altitudes, where the asymmetric model run shows a slight decrease in the SH speeds when $B_y < 0$, when compared to the symmetric case. An increase in the SH speeds, which may be expected when $B_y > 0$ is not clear however.

In figure 7.15, the dependence of the neutral wind speed on IMF sector is shown for the two simulations run during the second (September 2002) interval. Again the same strong dependence on B_z is seen in both models and the symmetric case shows slightly greater SH speeds. The B_y effect on the SH speeds is also observed much more strongly than the March 2001 simulation, where $B_y > 0$ causes an increase in the SH speeds for the asymmetric model when compared with the symmetric model. Conversely, $B_y < 0$ results in a decrease. The speeds in the NH for both the 2001 and 2002 simulations show little difference between the symmetric and asymmetric model.

The results are now compared with ground-based observations from FPIs in each hemisphere. In the NH, the Resolute Bay FPI is used, which is situated near the north magnetic dip pole at $+74.70°N, -94.90°E$, in northern Canada. In the SH, the Arrival Heights FPI at $-77.83°N, +166.66E$ is used, which is near to the south magnetic dip pole in Antarctica. FPI data from 630nm emissions are selected, from the same solar maximum in which the model runs were performed. Due to availability, the years 2002-2003 are chosen. The data, which are recorded close to the equinoxes (March, April, September and October), are binned according to the IMF sector, again using the OMNI data-set. The results are plotted for each hemisphere in figure 7.16. In each hemisphere the values from each of the asymmetric simulations at 240km are over-plotted where March 2001 is in blue and September 2002 in blue.



(a) Northern hemisphere

(b) Southern hemisphere

Figure 7.16: Dependence of average neutral wind speeds as a function of IMF sector for both CMAT2 simulations using asymmetric electric fields and FPI observations, for each hemisphere. The Resolute Bay FPI $(+74.70°N, -94.90°E)$ is used in the NH and the Arrival Heights FPI $(-77.83°N, +166.66E)$ in the SH. The simulations during March 2001 and September 2002 are now shown in blue and red, respectively.

The resulting values from FPIs exhibit little similarity between the model runs. The range of speeds are found to be significantly lower by the FPIs although there is a small increase in speed when B_z is negative, as is seen in the model. The FPIs show no clear B_y dependence. In the NH there is some increase in mean speed seen in sectors with opposite sign of B_y, where sectors 1, 2 and 3 have smaller values than sectors 7, 6 and 5 respectively. The range of values in each bin is large and each of these pairs of values are within one standard deviation of each other, suggesting the difference is not significant. In the SH, there is no clear difference in speeds as a function of B_y. The magnitude of speeds seen in each simulation is approximately equal, whilst the FPIs show lower values in the SH than the NH. The range in observed magnitudes, around 50 to $200ms^{-1}$, is much lower than is seen at the same altitude in CMAT2.

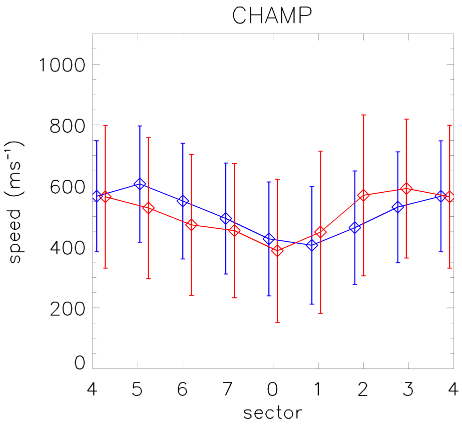


Figure 7.17: Neutral wind speeds binned according to IMF sector for CHAMP measurements. These plots are produced using tabulated data from Förster et al. (2008). Blue represents the NH and red the SH.

Next, figure 7.17 shows neutral wind speeds as measured by the CHAllenging Mini Payload Satellite (CHAMP) produced from tabulated data presented in Förster et al. (2008). The measurements from CHAMP are based on cross-track accelerations at $\sim$ 400km altitude, which is approximately the maximum height of CMAT2. The results from CHAMP show a slight difference between the NH and SH for opposite signs of B_y, which matches that seen in CMAT2 in the asymmetric simulations. The observations again show a large range in each bin, however, and the differences are again within one standard deviation. The absolute magnitude of the observed winds are slightly greater than those seen in the model at the same altitude.

7.4.2 Joule Heating

This section compares the height integrated Joule heating, Q_J, present in CMAT2 simulations with inferred values from ground based observations. This is achieved by combining measurements of Pedersen conductivity, σ_P, from IS radars at different magnetic latitudes with electric field measurements from SuperDARN. Here, data are selected from three locations covering different magnetic latitudes in the NH: Poker Flat ($\phi_m = 65.6°$), Sondrestrom ($\phi_m = 76.0°$) and Resolute Bay ($\phi_m = 82.8°$). The data are selected during the months of March and April in 2007 and 2008. These months are chosen to represent Spring equinox and the years are chosen due to the availability of data from all three radars during this period. Each conductivity measurement is multiplied by the corresponding E^2 value from SuperDARN, at that instant, and then binned according to IMF sector, using the corresponding values of B_y and B_z from the OMNI data set. The results are then further binned by MLT and altitude, to provide a two dimensional map of Joule heating power density as is shown in figure 7.18. The radars each occupy a fixed magnetic latitude, however the beam direction is variable. For the Sondrestrom radar, for example, the horizontal extent of the field of view in the E-region is ~ 90km (Thayer, 1998), which is less than the resolution of CMAT2 in either direction. As such, the horizontal variation of measurements is ignored.

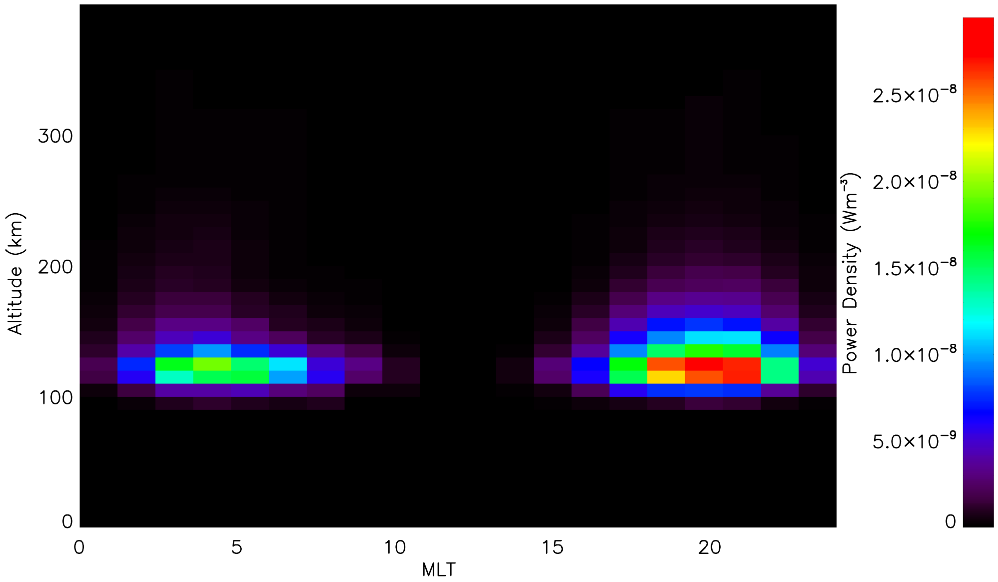


Figure 7.18: Joule heating power per unit volume calculated from σ_P and $\mathbf{E}$ measurements from Poker Flat IS radar and SuperDARN respectively. The plot represents data averaged over 10km bins of altitude and 18° bins of MLT for IMF sector 4.

The purpose of this binning procedure is to provide a means of comparison with the output of CMAT2 simulations. For each radar, a 2D map of Joule heating is produced for each of the eight IMF sectors, for example that of sector 4 is shown in figure 7.18. These are then summed over all altitudes to provide a measurement of height integrated Joule heating as a function of MLT for each sector, which are shown in figure 7.19. For comparison, the output from two simulations using SuperDARN and Foster's model to drive CMAT2 are plotted in figures 7.20 and 7.21. Figures 7.19, 7.20 and 7.21 are all plotted on the same (logarithmic) scale. It should be noted that, due to lack of IS radar measurements in the SH, it has not been possible to produce a map of observed heating in this hemisphere.

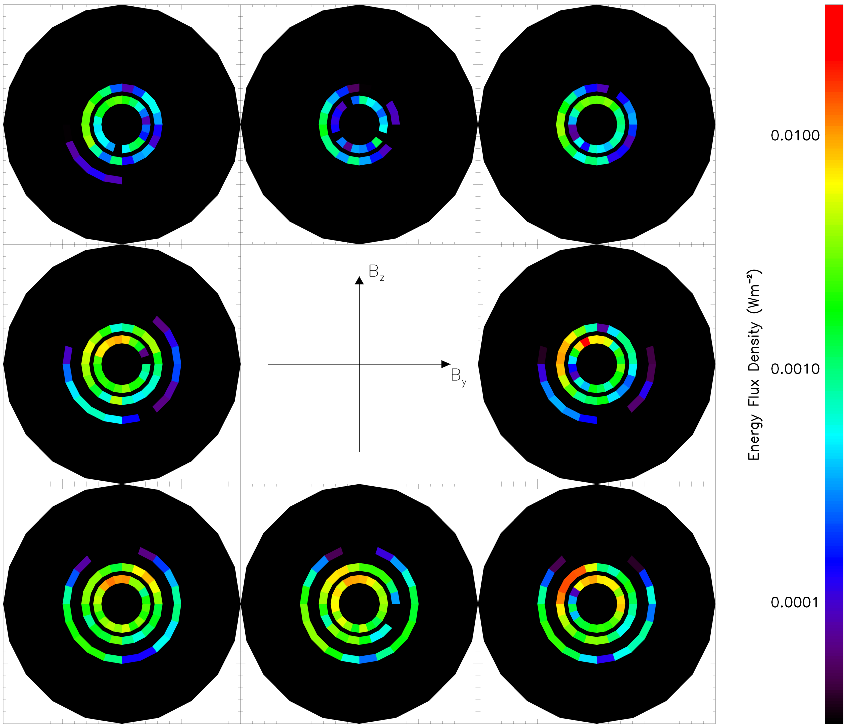


Figure 7.19: Average height integrated Joule heating determined from IS radar and SuperDARN measurements of σ_p and **E**. The results are binned according to sector and MLT, for magnetic latitudes corresponding to each of the three IS radar locations: Poker Flat ($\phi_m = 65.6°$), Sondrestrom ($\phi_m = 76.0°$) and Resolute Bay ($\phi_m = 82.8°$). The plots are each centred in magnetic coordinates with 12MLT to the top of each. The measurements are plotted within 2° of latitude of the radars position for clarity, although their actual coverage is less than this.

The plots in figure 7.19 show the height integrated Joule heating produced from IS radars and SuperDARN using the method described above. The concentric rings, from smallest to largest, correspond to the radar position of Resolute Bay, Sondrestrom and Poker Flat. Measurements at all three positions show a general increase in the magnitude of Joule heating as B_z becomes increasingly negative. This is to be expected, however, because the calculation of Joule heating involves an E^2 term from SuperDARN, which is known to increase with negative B_z, as is seen in figures 7.4, 7.6 and 7.8. Another feature that is apparent from figure 7.19 is that the magnitude of Joule heating is also increased when B_y is positive. This is particularly visible when B_z is strongly negative. The two innermost radars show heating that is greater on the day-side and peaks between around 12 and 18MLT. The Poker Flat radar at the lowest magnetic latitude shows more heating on the night-side, although the magnitude is much lower than for the other two radars.

Figure 7.20 shows the model output for the March 2001 simulation using asymmetric electric fields from SuperDARN. The Joule heating has been summed over each column of grid cells within the model, in the same way as the radar observations, to provide a value for height integrated Joule heating. The model exhibits the same trends with respect to B_y and B_z that are visible in the radar determined heating, with the peak occurring in sectors 2 (B_y strongly positive) and 3 (B_y positive and B_z negative). The simulated Joule heating peaks on the day-side, as was seen in the observations, although the model shows a peak of around 27.7mWm^{-3}, approximately twice that seen in the radars. The observations do have a limited coverage in latitude and a localised peak in heating could feasibly exist in a region of the atmosphere that is not observed by radars. The heating peak within the model occurs closer to 12MLT than the peak observed by radar, however the overall magnitudes seen in each are not dissimilar. In the SH, CMAT2 shows much lower overall heating. The B_z dependence is still apparent, however there appears to little difference between sectors with positive and negative B_y.

Lastly, figure 7.21 shows the model output from the March 2001 simulation using Foster's electric fields. These figures are plotted the same logarithmic scale as figures 7.20 and 7.19, which reveals the extent to which Joule heating is decreased in the Foster model, when compared with both the SuperDARN-driven simulation and the radar observed heating. A similar B_z dependence is seen in the magnitude of heating in the Foster-driven model, however there is no discernible B_y dependence. In this simulation there is no apparent difference in heating between each hemisphere.

Whilst the global plots in figures 7.19, 7.20 and 7.21 show the general distribution Joule heating, figures 7.22 and 7.23 allow a more direct comparison between the absolute values

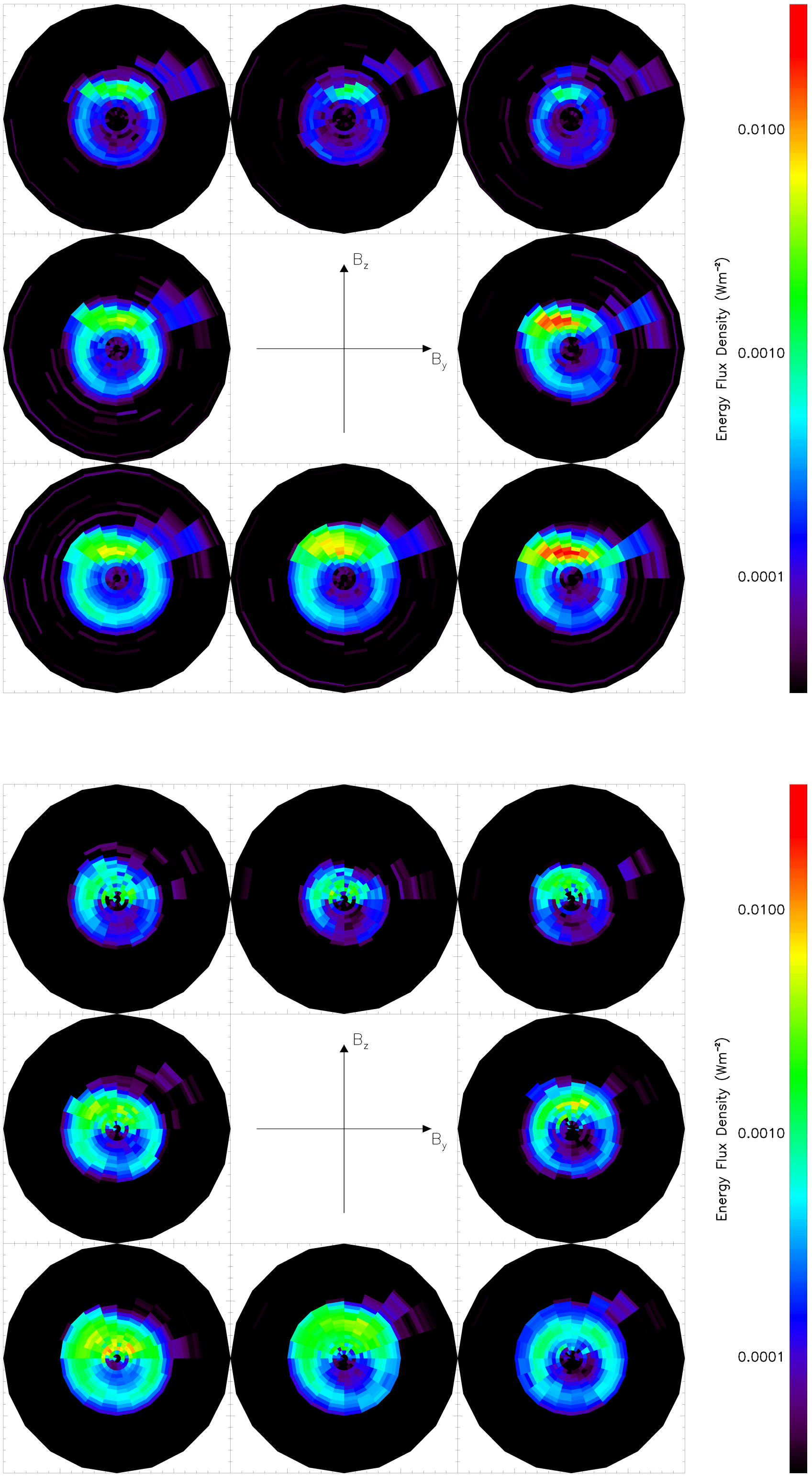


Figure 7.20: Average height integrated Joule heating from the March 2001 simulation using asymmetric electric fields and binned by IMF sector. The top plots correspond to the NH and the bottom plots show the SH. The plots are displayed in magnetic coordinates, with 12MLT to the top of each plot.

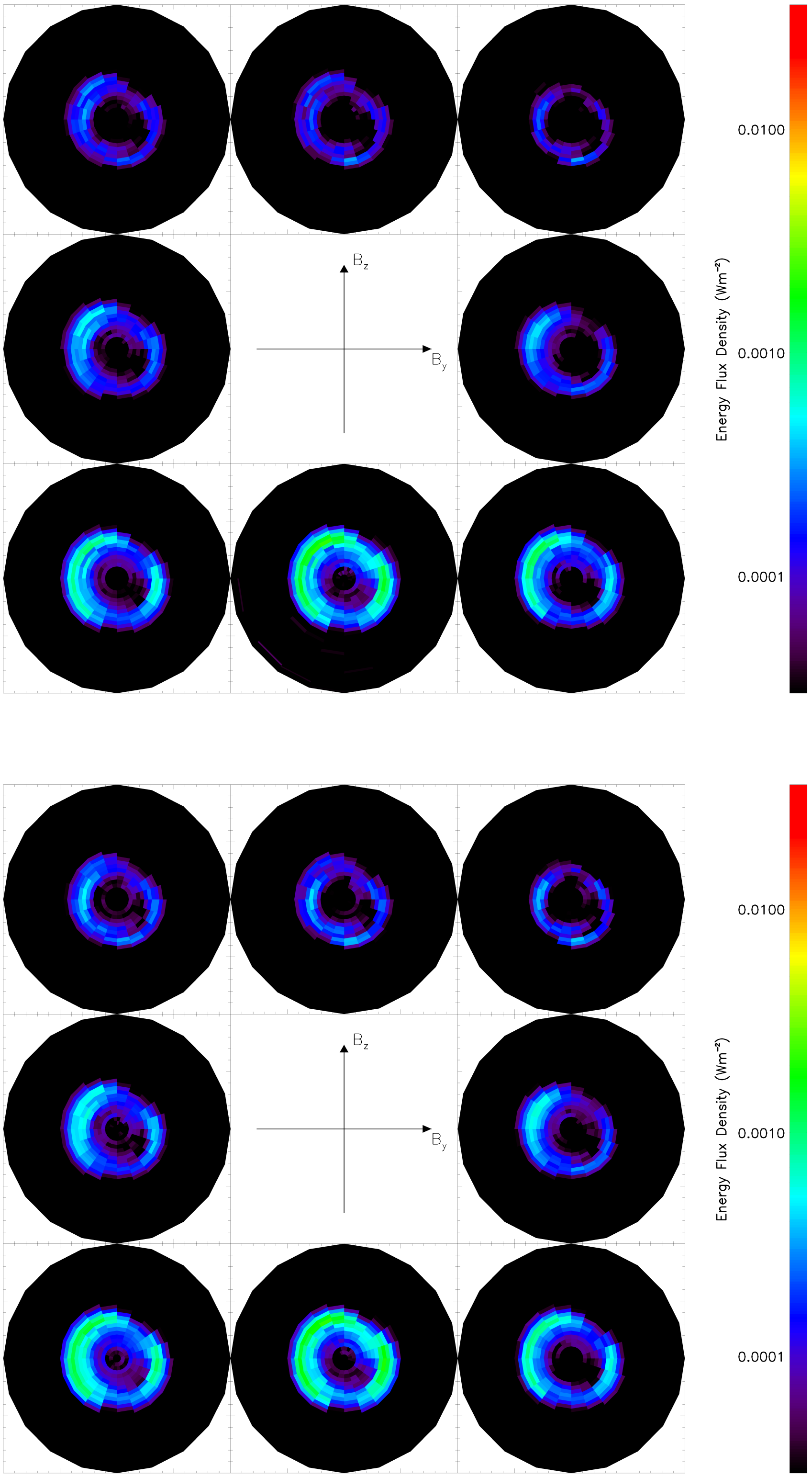
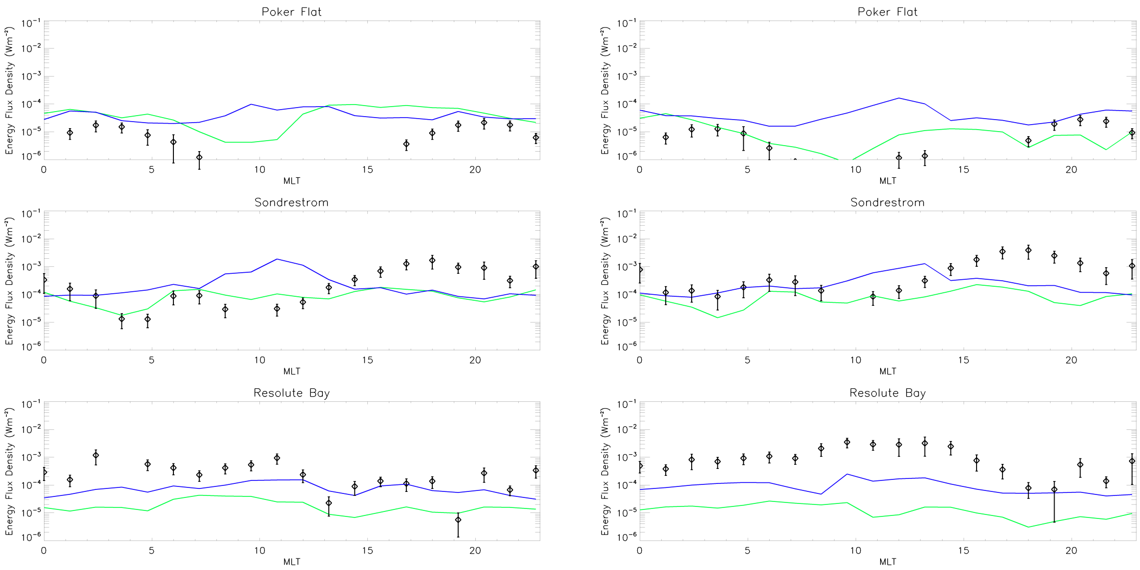
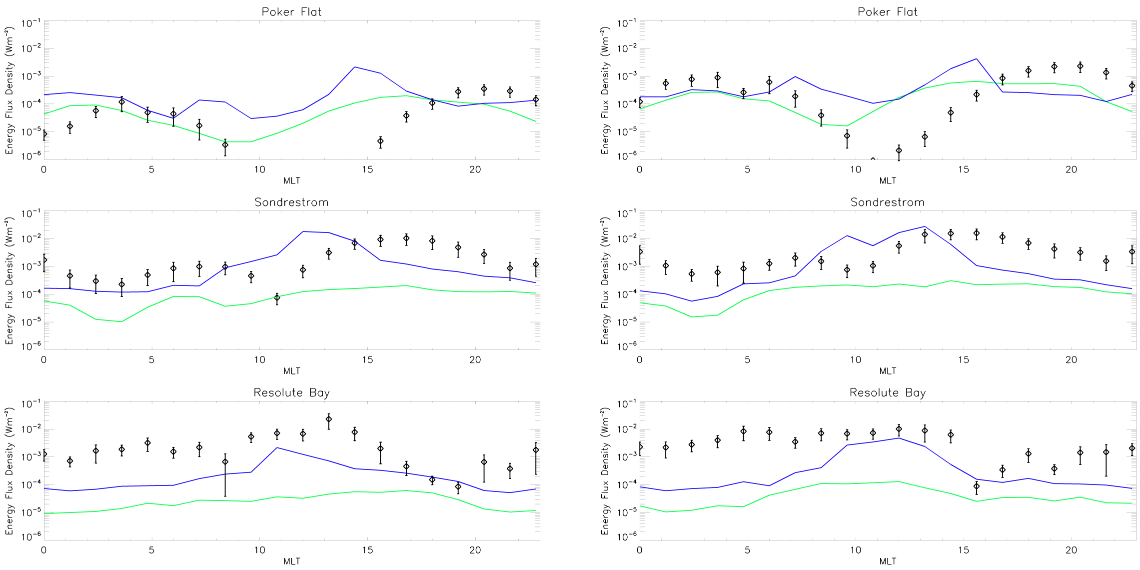


Figure 7.21: The same format as figure 7.20 but for the March 2001 simulation driven by Foster's electric field model.



(a) sector 0

(b) sector 1

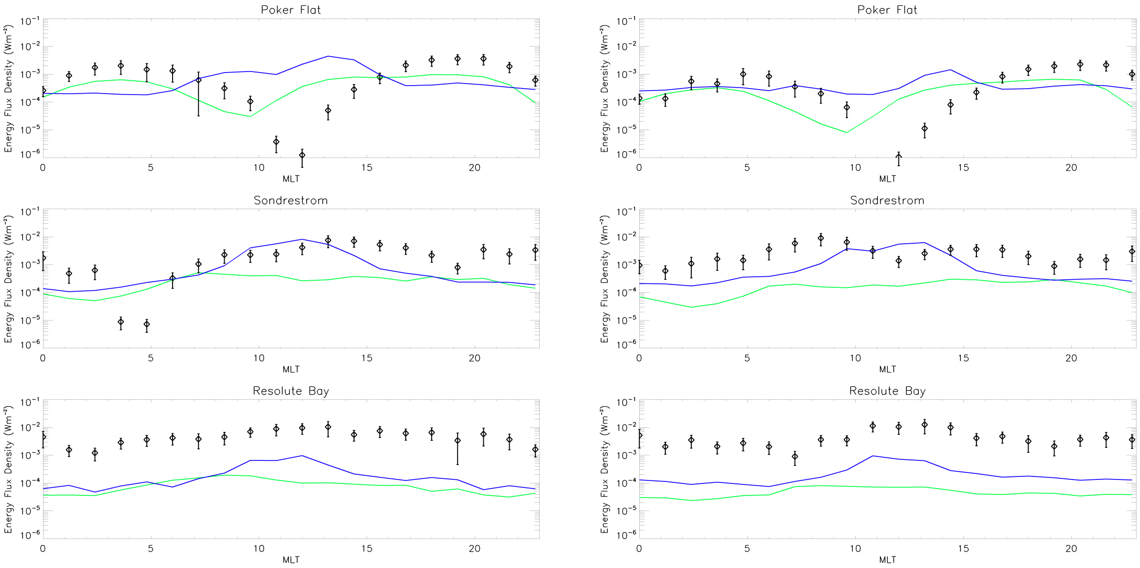


(c) sector 2

(d) sector 3

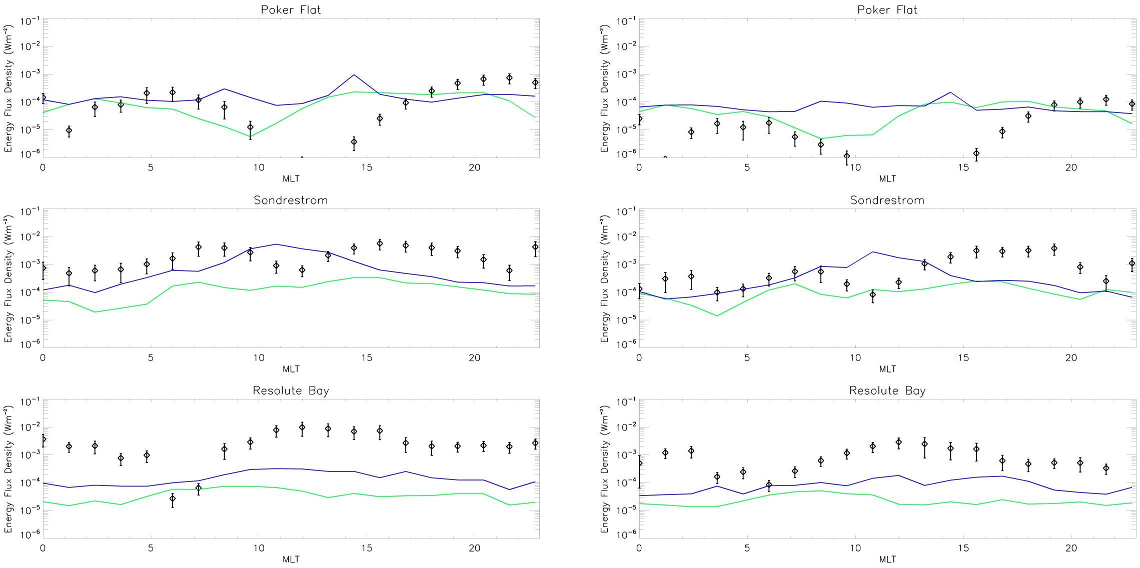
Figure 7.22: A comparison of height integrated Joule heating between model simulations and observations for IMF sectors 0 to 3 in the NH. The black data are the values determined from IS radar and SuperDARN measurements, averaged according to IMF sector and MLT. The values at the corresponding magnetic latitudes from model simulations using SuperDARN (blue) and Foster (green) electric fields are over-plotted for comparison.

seen in observations and simulations. There are eight sets of three plots, corresponding to the eight IMF sectors as measured at each of the three IS radar latitudes. The height integrated Joule heating is taken from each of the SuperDARN (blue) and Foster (green) model runs, at the latitudes corresponding to each radar, and plotted against the radar observations. All plots are set to the same scale on the $y-$axis.



(a) sector 4

(b) sector 5



(c) sector 6

(d) sector 7

Figure 7.23: The same format as figure 7.22 for IMF sectors 4 to 7.

At Poker Flat, $\phi_m = 65.6°$, there is a good agreement between the values seen in the Foster simulation and in the radar measurements. Both simulations, and the radar derived values, show that there is generally much less heating at this latitude when compared to the other two radar positions. The radars show some heating but it is only at all significant in sectors 3, 4 and 5, when the IMF is strongly southward. This heating peaks at about 4MLT and 19MLT, which is in agreement with the Foster model in terms of location. The magnitude of these peaks, however, are much lower in the Foster model, reaching about 1.0mWm^{-2} in sector 4, compared to 3.7mWm^{-2} in the radar observations and 4.5mWm^{-2}

using SuperDARN. So the SuperDARN driven simulations show a better agreement with the radar measurement in the magnitude of heating at this latitude, but produces the peak closer to 12MLT.

At Sondrestrom, $\phi_m = 76.0°$, the radar-derived heating shows a peak at around 15MLT in each sector, with a second peak at around 7MLT in 5 and 6 ($B_y < 0$). At this location, the heating within in Foster simulations is consistently and significantly lower than that seen in the radar observations. At the highest value, in sector 4, the Foster simulation reaches a value of less than 0.5mWm^{-2}. The location of the peak does, however, correspond to about 15MLT and there is a secondary peak at around 7MLT (when $B_y < 0$ in sectors 6 and 7), which is consistent with the radar observations. The magnitude of heating from the SuperDARN simulation is again much closer to the radar values than that of the Foster simulation. The strongest peak in heating occurs in sectors 2 and 3 in both the SuperDARN simulation ($\sim$18.2 and $\sim 27.7 mWm^{-2}$ respectively) and in the radar observations ($\sim$10.4 and $\sim$15.9mWm^{-2}). Again, however, the SuperDARN peaks occur closer to 12MLT than those observed by the radar.

Finally, at $\phi_m = 82.8°$, heating is derived from measurements at Resolute Bay. The range of values are of similar magnitude to those seen at Sondrestrom, except that they occur at different positions. The Resolute Bay radar generally shows a single peak at around 12MLT. The values are again greater when $B_z < 0$, but there is no clear dependence on B_y. At this latitude, both simulations show significantly lower heating than is observed. In the case of the SuperDARN simulation the highest value seen at this latitude is 4.8mWm^{-2} (sector 3) and in Foster it is less than 0.2mWm^{-2} (sector 4). The respective peak values measured by the radar in these sectors are 10.2 and 10.4mWm^{-2}. Although not clear in all the plots, both SuperDARN and Foster exhibit a single peak at approximately 11-12MLT, in each sector.

Figure 7.24 shows both the hemispheric power and global power from Joule heating, for each of the CMAT2 simulations, as a function of IMF sector, which are tabulated in 7.3. The dependence on B_z can be seen clearly, where both the Foster and SuperDARN models show greater heating in sectors where the IMF is strongly southward. The absolute values seen in the SuperDARN simulation are typically larger in the NH when $B_y > 0$, the greatest differences between the hemispheres can be seen in sectors 2 and 3. Conversely, sectors 5 and 6 show greater heating in the south, when $B_y < 0$. In the Foster simulation there is no discernible difference between the NH and SH and no B_y dependence is apparent from

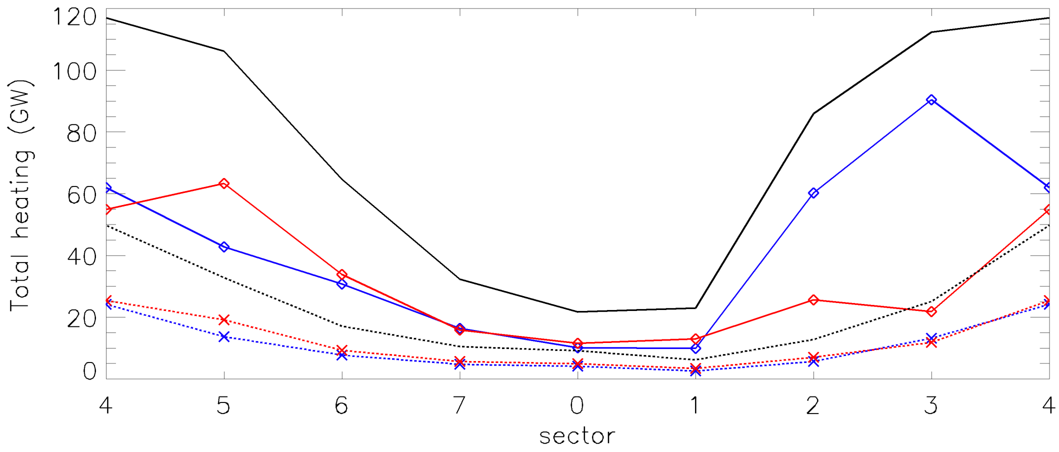


Figure 7.24: Total hemispheric Joule heating as a function of IMF sector from CMAT2 simulations using Foster and SuperDARN electric fields. The solid lines represent SuperDARN and the dotted lines Foster. NH is shown in blue, SH in red and the black lines are the global total.

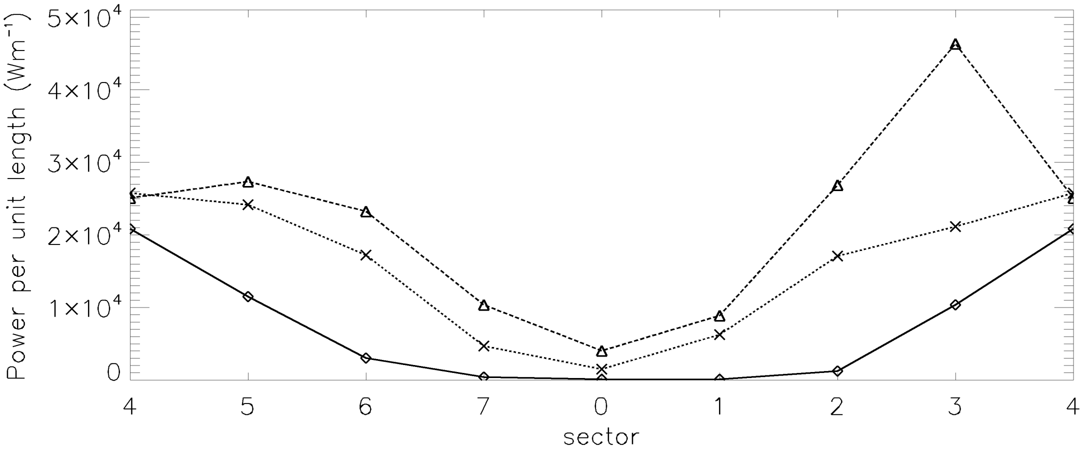


Figure 7.25: Joule heating determined by three IS radars, integrated over distance along magnetic local time. Poker Flat ($\phi_m = 65.6$) is represented by the solid line, Sondrestrom ($\phi_m = 76.0$) by the dashed line and Resolute Bay ($\phi_m = 82.8$) by the dotted line.

the plot. The total heating, shown in black, demonstrates how much lower the heating is in the Foster simulation compared to SuperDARN. At best the Foster model heating is only 43% that of SuperDARN (sector 4) and at worst it is 15% (sector 2).

The values of average NH Joule heating were determined by Knipp et al. (2005) for solar cycles 21-23, and were plotted in figure 2.14 in chapter 1. They determined values of 77±12GW for solar minimum, 112±18GW for solar maximum, 331±99GW for the top 5% most powerful events and 638±287GW for the top 1%. Of these, the solar maximum value

	SuperDARN			**Foster**		
sector	**NH**	**SH**	**total**	**NH**	**SH**	**total**
0	10.13	11.63	21.76	4.18	5.05	9.23
1	10.00	13.00	22.99	2.72	3.53	6.24
2	60.34	25.68	86.02	5.72	7.08	12.79
3	90.52	21.90	112.42	13.27	11.88	25.16
4	62.04	54.95	116.97	24.29	25.50	49.78
5	42.86	63.42	106.29	13.74	19.18	33.02
6	30.81	33.91	64.73	7.79	9.33	17.11
7	16.47	15.92	33.40	4.84	5.74	10.58

Table 7.3: Total hemispheric power in GW from Joule heating as a function of sector for SuperDARN and Foster driven simulations, corresponding to the values plotted in figure 7.24.

is the best comparison for the two CMAT2 simulations, which both took place during this part of the solar cycle. The Foster model shows heating significantly lower than this, with the highest value of only 49.78GW seen in sector 4. The peak value in the SuperDARN model of 90.52GW (also sector 3) is much closer to the value quoted by Knipp et al. (2005), although this is only true for strongly negative B_z values. For other less strongly southward IMF, less heating is seen in the SuperDARN driven model.

The limited latitudinal coverage of the radars mean it is not possible to produce a value for total hemispheric, or global, heating. However, figure 7.25 shows the radar values from figures 7.22 and 7.23 integrated along the distance covered by the radar as it moves in magnetic local time. This provides a value of heating per unit length, where length is in the north-south direction in magnetic coordinates. This is effectively a measure of relative total heating at a given latitude. The radars each show an increase in heating with negative B_z, however a B_y dependence is also seen in Sondrestrom. The value in sector 3 is much greater than that in sector 5, for opposite sign of B_y, and this difference is in the same direction as that seen in the SuperDARN simulation; an increase fro $B_y > 0$ in the NH. This is also the latitude at which the maximum heating (per unit length) is observed.

7.5 Discussion

7.5.1 Neutral Wind Speeds

CMAT2 has been used to simulate the response of the thermosphere to both hemispherically-symmetric and asymmetric electric fields over two five-day periods. These occurred during solar maximum, close to the March 2001 and September 2002 equinoxes. Figures 7.14 and 7.15 show the mean neutral wind speed above $|\phi_m| = 80°$ as a function of sector. In all simulations the neutral wind speeds under negative B_z conditions are increased in both hemispheres. This result is expected because of the of the greater electric fields that exist under these conditions. This is due to enhanced magnetic reconnection between the solar wind and the magnetosphere. In all simulations there was a slightly greater neutral wind speed in the SH, on average, which is found by Thayer and Killeen (1993) from DE2 observations. In the modelling study by Förster and Cnossen (2013), greater speeds are instead observed in the NH, which the authors attribute to the differing offset between their geographic and magnetic poles. Whilst CMAT2 uses a slightly different IGRF to that of CMIT in the Förster and Cnossen (2013) study, the greater SH speeds are thought to be due both to SuperDARN measuring stronger electric fields in that hemisphere and the IGRF having a larger magnitude, therefore producing a greater $\mathbf{E} \times \mathbf{B}$ drift. Under changing B_y conditions there was some difference in behaviour between the symmetric and asymmetric models, which is seen most clearly in the 2002 simulation (figure 7.15) at $240km$. The asymmetric simulation shows an enhancement in NH velocities when $B_y < 0$ and in SH velocities when $B_y > 0$ when compared to the symmetric simulation. Figures 7.4 to 7.9 show a small dependence of the cross cap potential on B_y, but it is in the adverse direction to account for this observed difference in speeds. The hemispheric difference in neutral speeds instead appears to be caused by the twisting effect of B_y on the ion convection pattern. This effect was used by Förster et al. (2012) to explain the same hemispheric difference seen in CHAMP observations (figure 7.17) from Förster et al. (2008). They claim that the clockwise rotation of the convection pattern for positive (negative) B_y in the NH (SH) causes the ion-neutral drag force to act at a greater angle to the pressure gradient force. The results in the net acceleration on the neutral atmosphere, from the momentum equation, to be reduced. For the opposite sign of B_y the two forces act in approximately the same direction, producing a greater net force. The effect is seen in the March 2001 simulation at $240km$ and it is clear in the September 2002 simulation at both 240 and 400km. Observations from two FPIs (figure 7.16), one close to each of the magnetic dip poles, fail to reproduce this hemispheric asymmetry seen in both CMAT2 simulations and CHAMP observations. The FPIs each provide a measurement from a single point on Earth, whilst CHAMP represents a global picture, averaged over many orbits. The latter therefore provides a better comparison to the model. The range of speeds seen in CHAMP at 400km altitude is slightly greater than that seen in CMAT2 at the same altitude, suggesting the model is underestimating speeds at this altitude. Conversely, the FPI measurements at 240km show much lower speeds than the model. The difference in magnitude between CMAT2 and CHAMP speeds is explained by the fact the SuperDARN is found to underestimate the magnitude of electrostatic potential (Cousins and Shepherd, 2010) when compared to other sources of measurement, as was discussed earlier in the chapter. One of these sources was the DCIM model (Papitashvili and Rich, 2002), which used measurements from DMSP to determine the cross cap potential in each hemisphere. The average values of cross cap potential from SuperDARN data are found to be 80% (NH) and 76% (SH) those of DCIM, whilst the average velocity measured by CMAT2 was found to be 76% (NH, 2001 simulation), 82% (SH, 2001 simulation), 75% (NH, 2002 simulation) and 83% (SH, 2002 simulation) that of the CHAMP measurement. It is likely then that the lower velocities seen at high altitudes in the model are the results of SuperDARN underestimating the polar electric field. This does not, however, affect the direction of the ion flow. Therefore the coupling between the pressure gradient and convection direction, which is thought to be responsible for the observed hemispheric differences is still present.

7.5.2 Joule Heating

Lastly, the power due to Joule heating in the model was studied. Height integrated Joule heating from IS radars at three magnetic latitudes in the NH was calculated using conductivity measurements from the radars combined with SuperDARN measurements of E^2, using the method described in section 7.4.2. Figure 7.19 shows the results of these observations, which are plotted on the same scale as CMAT2 simulations using both SuperDARN and Foster's model (figures 7.20 and 7.21) for 16-20 March 2001. These show the overall magnitude in heating to be significantly lower when CMAT2 is run using the Foster fields than is seen in the radar measurements. A modelling study by Rodger et al. (2001) using Foster's electric field model to drive CTIP with an hourly varying input was found to underestimate Joule heating by up to 65% (20% on average) when compared to values determined from the EISCAT IS radars. When used in CMAT2 with a three-hourly varying input, the Foster model typically produces heating several times lower than when SuperDARN is used instead. The heating values determined from the IS radars here are much closer to the SuperDARN simulation. However, both of these measurements rely on the SuperDARN electric fields, albeit from different time periods. Both of these measurements show the heating to peak in sectors 2 and 3, where the variability is also seen to peak (figure 7.8). This suggests that the cause of the increased heating is the extra variability. Comparing the SH heating in figure 7.20 to the electric field variability in figure 7.9 again shows a correlation between the sector where maximum heating and maximum electric field variability occur. However, the radar data in this hemisphere are too few to build up a map of heating, as with the NH in figure 7.19. The NH and SH heating in the Foster simulation (figure 7.21), both of which contain little variability, are significantly lower than the observed values and show no hemispheric difference. This would suggest that, as was found in section 6.2, that the electric field variability does strongly increase Joule heating. The maximum standard deviation seen in figures 7.8 and 7.9 is less than 15kV, whilst the maximum cross-cap potential difference is over 60kV. This shows that, at most, that σ is approximately one quarter the value of e_m. Codrescu et al. (2000) have claimed that these two quantities contribute equally to heating, whilst the results here suggest σ may in fact be more significant. Their claim was based on the unjustified assumption that electric field variability follows a normal distribution, which figures 7.1, 7.2 and 7.3 show is not the case.

Figures 7.22 and 7.23 show a direct comparison between simulations and observations for the three magnetic latitudes at which the IS radars are located. The limited latitudinal coverage of the three radars used in this study, when compared to a whole atmosphere model, means that comparisons may only be made at specific latitudes, rather than globally. For example, the location in which the peak heating occurs in the SuperDARN model in figure 7.20 (sector 3) is at a magnetic latitude between the Sondrestrom and Resolute Bay radars. This means that, were it present in the true atmosphere, it would be not be observed by either of these radars. In figures 7.22 and 7.23 above, the locations (in MLT) and strengths of the peaks in height integrated Joule heating were examined. Whilst the heating magnitudes were generally more similar to the SuperDARN simulation, the locations were found to be closer to the Foster simulation. It should be noted that both the SuperDARN and Foster driven simulations use the same particle precipitation pattern. This uses K_p as a proxy and therefore matches the Foster electric field model well. Due to the lack of a better precipitation model, this must also currently be used with SuperDARN electric fields. The poor correlation between K_p and SuperDARN electric field strength was demonstrated in the previous chapter, for example in figure 6.2. The precipitation model contains a total ionisation rate, which in turn determines the ionisation rates of the major constituents, as is described in section 4.2.4. The distribution of this ionisation controls the distribution of conductivity in the model, and therefore influences Joule heating via the equation $q_J = \sigma_p E^2$. The locations of the peaks in MLT in the Foster simulation are generally in agreement with the radars at all three magnetic latitudes. In the SuperDARN simulation this is not the case and the peaks typically occur closer to 12MLT, including the cases where there are two observed peaks at $\phi_m = 65.6$ (Poker Flat) in sectors 2, 3 and 6. Because the Foster electric field model has a matching precipitation pattern, one would expect that it will accurately model the location of Joule heating because the values of σ_P and E^2 are compatible, at a given location. However, the absolute magnitude of the heating is much lower due to the lack of variability. In the SuperDARN simulations the conductivity as a function of position, compared to the electric field, does not necessarily agree and the model produces Joule heating with an incorrect distribution. Instead, the location of the auroral oval is observed by Imber et al. (2013) to correlate with SuperDARN's HMB and so an improved precipitation model, to match the SuperDARN observations, is needed by CMAT2.

Finally, the total hemispheric Joule heating from each simulation was compared in figure 7.24. As is apparent from the plots in figures 7.20 and 7.21, the total heating in the SuperDARN driven simulation is consistently greater than that which used the Foster fields. Additionally there is little difference between the NH and SH in the Foster simulation, whilst in the case of SuperDARN the NH is greater. Both of these features again suggest the sectors containing a more variable electric field in figures 7.8 and 7.9 produce more heating, whilst the static Foster field produces little. NH Heating power in the SuperDARN simulation is more than double that in the SH for sectors 2 and 3 when $B_y > 0$. The opposite difference between north and south is observed when $B_y < 0$ in sectors 5 and 6, but is found to be much smaller. When applying his electric field model to estimating Joule heating, Weimer (2005) also produced NH values that were very slightly greater when $B_y > 0$, which were not explained by the author. The two electric field regions differ between the hemispheres in both their magnetic field strength and their geographic-magnetic polar offset. However, neither of these features should be expected to affect the

Joule heating. This is supported by the fact that the Foster driven simulation, which uses the same magnetic field as SuperDARN, shows little hemispheric heating difference (figure 7.24). A difference, which is relevant only in the SuperDARN simulation, is the coverage of the radar network in each hemisphere. There are fewer radars in the SH, which typically means less back-scatter is observed at equinox, compared to the NH. This means the potential distribution in the SH is more dependent on the IMF-dependent empirical model. Because of this, SuperDARN electric fields, when binned according to IMF-sector, will exhibit less variability than real-time radar measurements. This is exactly what is shown in figures 7.8 and 7.9. Because the variability has a strong influence on heating, as has been demonstrated throughout this chapter, then the lack of SuperDARN observations in the SH will account for the lower Joule heating seen in the model.

The heating in figure 7.24 was compared to the power per unit length from the three radars as a function of sector (figure 7.25). The Sondrestrom radar, at the magnetic latitude where the greatest heating occurs, shows an increase in power when $B_y > 0$, in agreement with the NH in the simulation. These observations are also reliant on SuperDARN measurements of the electric field to calculate heating, so this difference would again suggest variability may be the source of increased heating.

When compared to the total northern-hemispheric Joule heating published by Knipp et al. (2005) at solar maximum (112 ± 18GW), the SuperDARN simulation was found to be somewhat similar (90.52GW) for sector 4, where most heating occurs, but typically lower in other sectors, with an average value of 44.48GW. The neutral velocities within CMAT2 were also found to be lower when compared to measured values from CHAMP and it was suggested that this may be due to SuperDARN underestimating the electric field strengths. Due to the dependence of Joule heating on E^2, it should be more sensitive to the amount by which SuperDARN underestimates the cross cap potential (80% in the NH). Simply squaring this value means we might expect the model to underestimate heating by 64% in this hemisphere. However, this effect is still not great enough to completely account for the difference between the heating observed in CMAT2 and the value presented by Knipp et al. (2005).

Summary

The effects of asymmetry between the electric fields in each hemisphere have been considered. CMAT2 is driven for two five day periods in which there was reasonably good coverage of both hemispheres in SuperDARN. These are $16^{th} - 20^{th}$ March 2001 and $14^{th} - 18^{th}$ September 2002, during solar maximum. The maximum and minimum electrostatic potential values in the respective dawn and dusk cells (figure 7.11) follow the IMF dependence in agreement with previous studies. The effect on the neutral winds in CMAT2, however, is that the SH experiences stronger neutral winds than the NH when $B_y > 0$ and the opposite is true when $B_y < 0$ (figure 7.14). This result is not strongly supported by observational evidence from FPIs, but is in agreement with CHAMP observations (figure 7.17), which is attributed to B_y positive (negative) producing a large dusk convection cell in the NH (SH). This acts in a direction that opposes, to some extent, the pressure gradient flow over the pole and reduces the overall neutral wind speed. When the sign of B_y is opposite, the dusk cell is reduced and the direction of the $\mathbf{E} \times \mathbf{B}$ complements the pressure gradient. This results in the hemispheric dependence of neutral speeds on B_y, which is seen in CMAT2.

The dependence of hemispheric Joule heating on electric fields was studied. Height integrated Joule heating from two CMAT2 simulations was calculated in each hemisphere. Both simulations were run for 16-20 March 2001, with one using asymmetric SuperDARN electric fields in each hemisphere and the other using Foster's model. The results were again binned according to IMF sector, in which the SuperDARN model showed consistently greater heating than Foster. In both models there was an increase in heating for negative B_z (figure 7.24) conditions, however the SuperDARN model also showed a B_y dependence in the NH. A general increase in variability as a function of sector is seen in SuperDARN in both hemispheres (figures 7.8 and 7.9). This increase in variability occurs in the same sectors as those which correspond to the greatest heating in the SuperDARN simulations. The Foster simulations, which possess no variability show much lower heating. When compared to radar observations (figures 7.19, 7.22 and 7.23), the Foster driven simulation was more consistent in the location at which heating occurred, which is thought be due to CMAT2 using a particle precipitation model designed to match Foster's electric fields. Conversely, the SuperDARN driven simulation was more consistent with the overall magnitude of the heating. The magnitudes of the Foster and SuperDARN electric fields are similar (e.g. figures 6.2 and 6.4), suggesting the increase in heating is a result of the variability in SuperDARN. The non-Gaussian distributions of cross-cap potential, for strong IMF conditions, in figures 7.1, 7.2 and 7.3 also suggest that the variability may contribute more to the Joule heating than from a normal distribution.

Both the velocities and heating magnitudes are underestimated by CMAT2 when compared to other sources of measurement. CHAMP velocities at 400km are approximately 20% to 35% larger than those seen in CMAT2, whilst hemispheric Joule heating is several times lower than that measured by Knipp et al. (2005) for solar maximum. SuperDARN is seen to underestimate the cross-cap potentials, which have values 80% (NH) and 76% (SH) those of DMSP measurements. The respective E and E^2 dependence of neutral winds and Joule heating hints that this underestimation contributes strongly to the low values seen in CMAT2. The heating estimates by IS radars used here also rely on SuperDARN, which means these are also underestimated by the same amount as the model.

This page was intentionally left blank

Chapter 8

A New IMF-Driven Electric Field Model for CMAT2

The work presented so far in this thesis has been sub-divided into two areas of research; one of observing solar wind and CME properties and one of modelling the terrestrial response to these phenomena. It was the original objective of this thesis to interpret the HI data in a way that could be used as an input to the CMAT2 ionosphere, and to develop CMAT2 in such a way that it could be driven by this input. This was with the intention of both improving the forecast times of geoeffective IMF parameters beyond the one-hour currently available from L1 spacecraft and of developing our ability to model their effect on the Earth. Of course, these ambitions are found to be quite beyond the scope of a single PhD. Chapter 5 has demonstrated that it is possible to infer some limited information about Earth-directed solar wind density and velocity from HI, however it is certainly not sufficient to drive the model. The conclusions of chapters 6 and 7 highlight that much more information is needed to drive neutral wind speeds and Joule heating, particularly the inclusion of IMF B_y and B_z components, and with a high temporal resolution in order to incorporate the effects of variability. A HI-driven CMAT2 is therefore not possible at this stage and it is instead the aim of this final chapter is to demonstrate the potential to drive CMAT2 using a new electric field model, which takes account of the findings in previous chapters. Existing GCMs generally use IMF input, either by Weimer's empirical model or via a coupled MHD magnetosphere module, however these are each known to possess weaknesses. It is the purpose of this new electric field model to confront these weaknesses by taking a new approach to quantifying the electric fields from SuperDARN in terms of IMF parameters.

It is a problem common to all GCMs that it is difficult to simulate the correct neutral wind speeds and Joule heating rates that result from interaction between the ionosphere and neutral atmosphere. The Global Ionosphere Thermosphere Model (GITM) (Ridley et al., 2006), which uses Weimer's electric field model, is found to underestimate Joule heating. Deng and Ridley (2007) attribute this to three sources, which are common to all GCMs; **E** field variability, grid resolution and vertical velocity. The first of these refers to temporal resolution of the electric field. The authors claim that a model which updates its electric field at least every three minutes is sufficient to produce correct temporal variability. The second source refers to spatial resolution of the electric field. By increasing the GITM resolution from $5° \times 5°$ to $2.5° \times 0.3125°$, Yiğit and Ridley (2011) find that Joule heating can increase by up to 50%. Lastly, the difference in the vertical component of ion and neutral velocities is shown to increase Joule heating if it is properly included in the model, however this increase is less significant than the other two sources.

The spatial variability is further divided into two sources by Matsuo and Richmond (2008), who identify variability on a scale less than, and greater than, model resolution as contributions Joule heating. These are coined *subgrid-scale* and *resolved-scale* respectively and are studied using TIE-GCM. The subgrid variability is found to produce an increase in heating, particularly in the winter hemisphere, and its absence is therefore a further cause of low Joule heating values in simulations. To account for this lack of subgrid variability, Emery et al. (1999) attempt to introduce a correction factor to the Joule heating within the model. However, Shiokawa et al. (2007) conclude that it is difficult to tune the value of this factor accurately because the small scale variability is both time-dependent and spatially varying. The CMIT model is a development TIE-GCM, which uses the LFM magnetosphere model to produce high-latitude electric fields, rather than relying on Weimer's model. Many studies using CMIT have found that it tends to overestimate the cross-cap potential, particularly during strong IMF conditions. Förster and Cnossen (2013) identify this as a source neutral wind overestimation in the model, whilst Burns et al. (2008), Wang et al. (2008) and Wiltberger et al. (2012) find that energy input is also overestimated, particularly during storm periods. Burns et al. (2008) suggest that the overestimation of cross-cap potential is the failure of CMIT to include saturation effects identified by Siscoe et al. (2002), which act to reduce its value.

Like CMAT2, the TIE-GCM and CMIT models use a pressure-based vertical coordinate system and are unable to account for the vertical velocity contribution to Joule heating that is found in the altitude-based GITM model. However, this is less important than temporal and spatial variability. The temporal resolution of all four models is sufficient to account for the temporal variability of less than three minutes identified by Deng and Ridley (2007). The spatial variability of each model is dependent on the electric field input. The LFM magnetosphere (CMIT) is found to overestimate the electric field magnitude compared to both observations and other models and is unable to produce the B_y dependence of the dawn and dusk cell orientation (Zhang et al., 2011). The Weimer model used by GITM and TIE-GCM does not suffer either of these problems but does lack spatial variability. Because it is based on averaged electrostatic potential measurements, the electric fields appear overly smoothed in longitude and latitude compared to observations from SuperDARN. Additionally it assumes a one-to-one correspondence between the solar wind parameters and potential distribution, which is not entirely accurate. The aim of this chapter is to develop a new empirical electric field model that quantifies this spatial variability and can be driven using the IMF as an input. This model is described in the following section and is then used to drive CMAT2.

8.1 Quantifying the Variability within the SuperDARN Electric Fields

An earlier attempt to include electric field variability in CMAT2[1] based on the findings of Codrescu et al. (2000), was pursued by Harris (2001) to account for the lack of Joule heating resulting from Foster's electric field model. This was achieved by adding a Gaussian random variable of standard deviation 0.01Vm^{-1} to the existing electrostatic potential values, if they exceeded 0.001Vm^{-1},

$$\mathbf{E} = \mathbf{E}_{Foster} + \sigma z \tag{8.1}$$

where $\sigma = 0.01$, z is a Gaussian random variable, $\mathbf{E}_{Foster}$ the original electric field vector

---

[1] This inclusion was dropped from CMAT2 and is not present in any versions of the model used in this thesis.

and **E** the new electric field vector. There are two noticeable weaknesses to this method; firstly that the value of σ is chosen arbitrarily and secondly that only the temporal variability, and not the spatial variability, is improved. Every point in the electrostatic potential distribution is increased by the same factor, σz, and so the direction and overall shape of the electric field do not change, only the magnitude. As, for example, Deng and Ridley (2007) show, in a model where the shape of the convection pattern remains constant, the neutral velocities tend toward the ion velocities and Joule heating values are found to be lower. An attempt is made here to address both of these weaknesses by developing a new model based on SuperDARN observations. This model quantifies the variability in the electrostatic potential maps, not only in terms of the standard deviation at each grid-point, but in terms of the covariance between grid-points. This provides CMAT2 with a variable electric field model that is based on, and possesses the same covariance properties as, the ionosphere as measured by SuperDARN.

As with the empirical models of Foster and Weimer, it is necessary to choose a proxy by which the data is binned and by which the appropriate electrostatic potential map may be called during a model simulation. Based on the conclusions of chapter 6, and on established wisdom, the IMF vector is the best choice. The data are therefore binned into the eight IMF sectors and three levels of IMF strength, according to $|\mathbf{B}| < 5.0$nT, 5.0nT$\leq |\mathbf{B}| < 7.5$nT and 7.5nT$\leq |\mathbf{B}|$, where these values have been chosen to produce an approximately equal number of counts in each bin. The eleven years of SuperDARN data from each hemisphere are selected and any data sets that were produced using fewer than 200 measurements are discarded. The remaining data sets are sorted into one of each 24 bins, based on the current IMF strength and orientation from the OMNI high resolution data set. For each SuperDARN data set a map of electrostatic potential is produced over a grid of 2° magnetic latitude and 18° magnetic longitude, covering $45^\circ \leq |\phi_m| \leq 90^\circ$ in each hemisphere. This corresponds to an array containing $N = 460$ elements, each representing a grid-point in magnetic latitude and longitude.

8.1.1 Mathematical Derivation of Electric Field Variability

The variability in each grid-point may then be quantified in terms of the covariance between that grid-point and every grid-point, including itself. This is represented by the $N \times N$ covariance matrix, $\mathbf{\Sigma}$, where the element i, j represents the covariance between grid cell i and j,

$$\mathbf{\Sigma}_{ij} = \mathrm{cov}(x_i, x_j) = \frac{1}{T}\sum_{i=1}^{T}(x_i - \mu_i)(x_j - \mu_j) \tag{8.2}$$

μ is the mean value in each grid cell and T is the total number of data sets contributing to a given IMF bin. The N diagonal elements of Σ are then the variance in each bin. The covariance matrix may be used in a similar way to equation (8.1) to produce a random electrostatic potential, except that now there is a physical basis for the variability and it contains spatial variability with the same covariance properties as the SuperDARN observations,

$$\mathbf{\Phi} = \boldsymbol{\mu} + \mathbf{Az} \tag{8.3}$$

where A is a matrix that satisfies

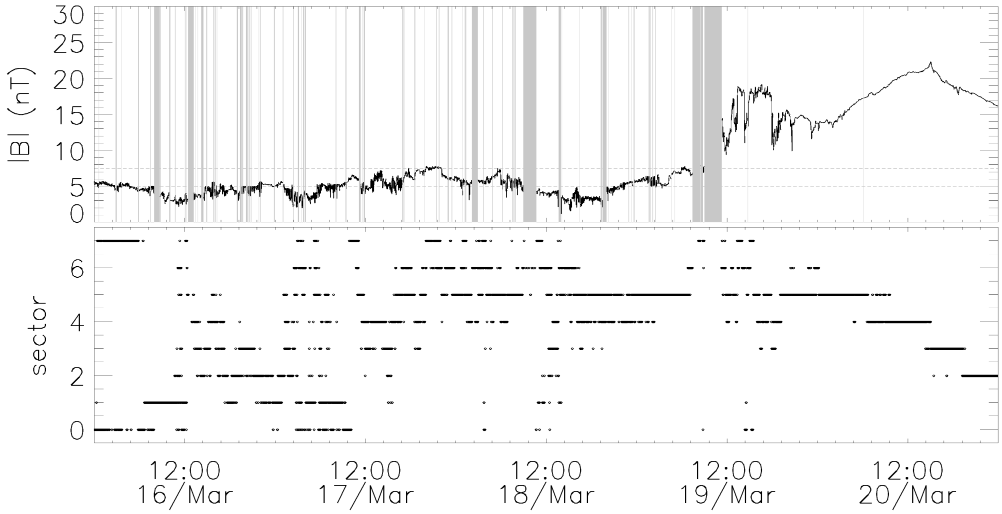
$$\mathbf{AA}^{\mathsf{T}} = \mathbf{\Sigma} \tag{8.4}$$

$\mathbf{z}$ is now an array of N independent Gaussian random variables and $\mathbf{A}$ may be thought of as the square root of $\mathbf{\Sigma}$, which may be determined using an iterative process called *LDLT decomposition*, (Appendix B). Because the covariance matrix is based on SuperDARN data with a two-minute cadence, these random fields are read in to the model every other 60 second time-step.

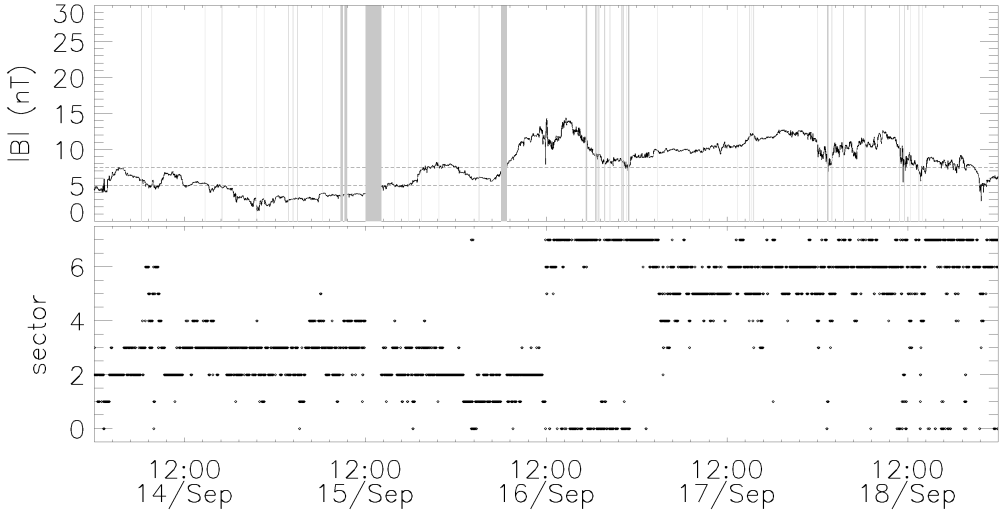
8.1.2 Simulations

As a means to test the effectiveness of this new model, particularly the effect of the added variability on neutral heating and winds, it may be run in two modes. The first of which contains the random variability, as described above using equation (8.3). The second mode neglects the $\mathbf{Az}$ term in equation (8.3) and so just produces the mean electric field conditions, similar to the Weimer model. For clarity, these are referred to as the *variable* and *mean* electric field models, respectively.

CMAT2 is run for two five-day periods using both the mean and variable models. The results from the previous model runs, which were driven directly by SuperDARN data, are



(a) March 2001



(b) September 2002

Figure 8.1: The strength and sector of the IMF as measured by wind for each simulation interval. The horizontal dashed lines in the |**B**| plots indicate the three levels of IMF strength in to which the SuperDARN data are binned. Grey regions represent a lack of data from the spacecraft, upon which the last 'good' data point is read in to the model instead.

also included for comparison. As such, the same two simulation intervals as in the previous chapter (16-20 March 2001 and 14-18 September 2002) are used. The mean and variable models are each driven using the values of B_y and B_z from the OMNI data-set over these five day periods. The aim of using these three different versions of the electric field model is to provide a clear demonstration of the effect of the added spatial variability on neutral winds and Joule heating.

Figure 8.1 shows the state of the IMF strength and orientation for the two five-day simulation intervals from the OMNI data-set. In the first (March 2001) simulation, the IMF strength remains fairly moderate, below 7.5nT, for the first 3.5 days until 1200UT on 19^{th} March. After this time it increases significantly, peaking at close to 25nT at 1200UT on 20^{th}. The IMF sector varies quite strongly during the first four days of the simulation until the end of 19^{th} March. After this, during the final 24 hours, the IMF B_z takes on quite strongly negative values between sectors 2 and 5. This period coincides with the period of very strong IMF magnitude. The second simulation interval (September 2002) shows a moderate IMF strength for the first 2.5 days, up to around 1200UT on 16^{th} September. After this, during the final 2.5 days it persistently exceeds 7.5nT. The IMF B_y component shows generally positive values during the period of weaker IMF strength in the first 2.5 days. In the final 2.5 days of the simulation, negative B_y values are seen, which coincide with the stronger IMF magnitude.

8.2 CMAT2 Results

Figure 8.2 shows an example of the three electric field models used to drive CMAT2; direct input from SuperDARN and both the IMF-dependent mean and variable electric field models. The five columns represent consecutive two-minute time steps that have been chosen during the strongly southward IMF in 2001 simulation interval. Over this short timescale there is no significant change in the IMF values and so the mean and variable models are using data from the same IMF bin; $|\mathbf{B}| > 7.5$nT and sector 4. As a result the mean electric field model remains constant during this period, however the variable model shows some change in strength and spatial distribution. The SuperDARN data shows a similar variability on this timescale, however the exact strength and spatial distribution of the variable model does not match the SuperDARN model because it is created artificially, at random. Over much longer periods, the mean and variable models would be expected to produce the same average electrostatic potential distribution as SuperDARN and the variable model would produce the same covariance properties too.

This figure is intended as an example to illustrate the different electric field models. Although this only shows ten minutes, it can be seen from figure 8.1 that the values of B_y and B_z can remain approximately constant, sometimes for hours. For example from 0938UT to 1420UT on 20^{th} March 2001, the IMF sector (4) does not change and it's strength varies little (between 19.3nT and 21.3nT). Therefore an mean electric field model

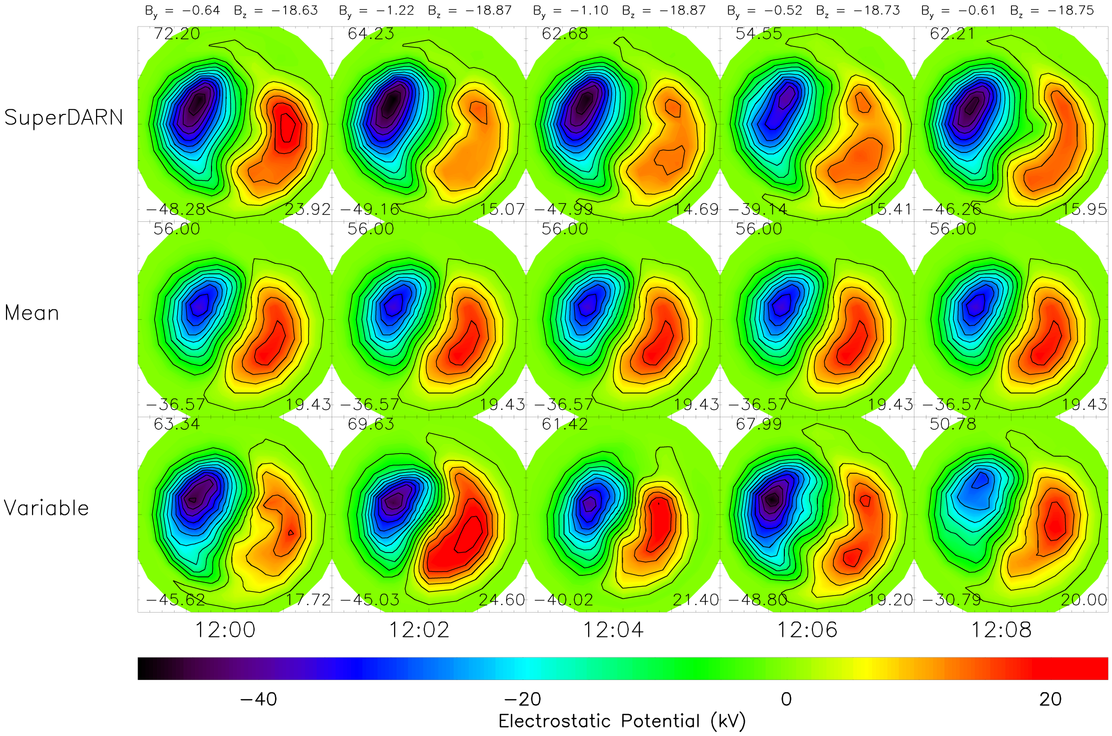


Figure 8.2: Electrostatic potential in northern hemisphere, with 12MLT to the top of the page. The top row shows five successive observations from SuperDARN at 1200–1208UT, 20^{th} March 2001. The IMF values from the OMNI data set throughout this period (displayed above) are used to produce the second and third rows using the respective mean and variable electric field models. In each plot the cross-cap potential is displayed in the top left and the minimum and maximum potentials are shown in the bottom left and right, respectively.

will remain unchanged for this entire period, whilst the real electrostatic potential is seen to change in strength and distribution during this period. In the variable model, at each 2-minute time step, random numbers are generated and then used to produce the variability of the potential. This means that there is no dependence of the potential distribution in a given time-step on its value in the previous time-step. The variability seen in the SuperDARN data over this period suggests that this dependence does in fact exist, where the distributions vary smoothly over time.

8.2.1 Electric Fields

The electric field strength, from each of the SuperDARN, mean and variable models, over the five day ($16 - 20^{th}$ March 2001) simulation is shown in figure 8.3. For each hemisphere, the average electric field magnitude from grid-points within $|\phi_m| > 80°$ is shown. The value derived from SuperDARN observations (top panel) is seen to begin relatively weakly

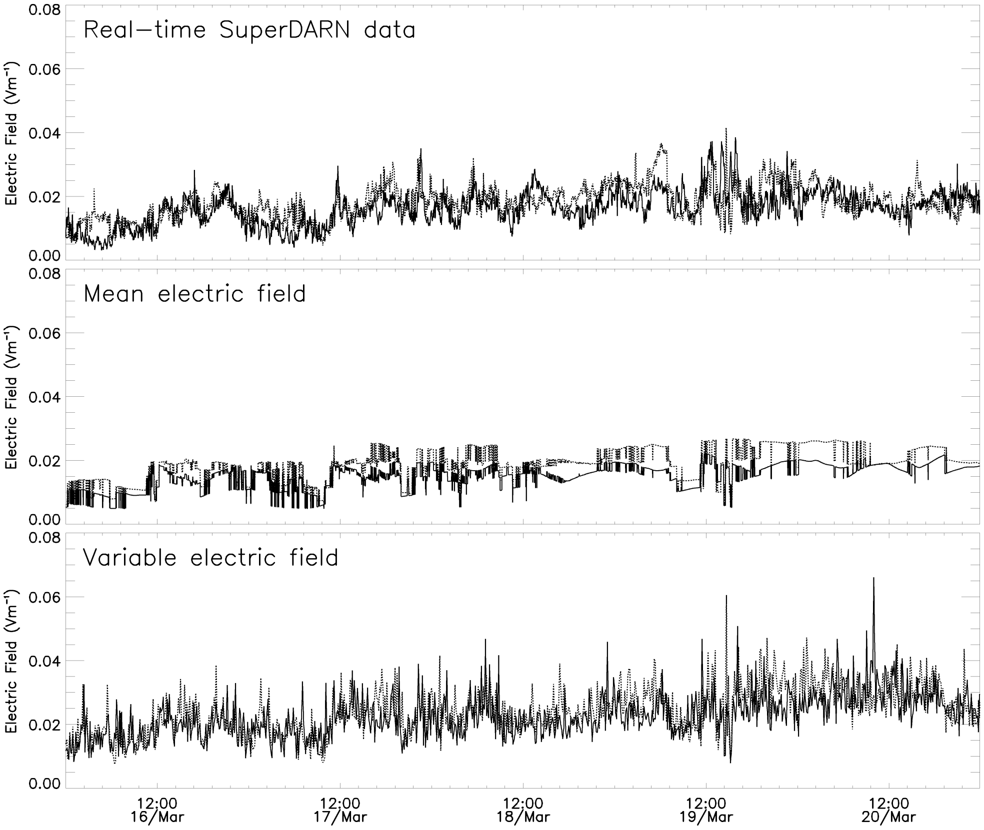


Figure 8.3: The mean electric field strength for the March 2001 simulation interval for SuperDARN and both the mean and variable electric field models. The mean is taken within the region $|\phi_m| > 80°$, with the NH and SH represented by the solid and dashed lines, respectively.

at around 0.01Vm^{-1} and increase throughout the simulation, peaking at around 0.04Vm^{-1} at 1200UT on 19^{th} March. During the final 24-hours of the simulation, when strong and southward IMF is observed by WIND, the field strength in both hemispheres remains just above 0.02Vm^{-1}, which is weaker than we may expect from these conditions. The mean electric field model follows the same trend, beginning at around 0.01Vm^{-1} and increasing to around 0.02Vm^{-1}. The variable model also increases over the same timescale but appears to show overall stronger values than the SuperDARN data, peaking at 0.06Vm^{-1} at 1000UT on 20^{th} March. The difference between the structure of the mean and variable electric fields is very apparent. The step changes seen in the mean model correspond to the IMF changing between sectors or strength-levels used by the model. At some points, particularly during the final 72-hours, the mean field is seen to be very static. This is caused by the same IMF strength and orientation persisting for a prolonged time. The time-variability in the variable electric field model resembles the SuperDARN model more closely, however the former appears to fluctuate more rapidly. This may be caused by the lack of correlation between successive time-steps in the variable model, which is seen in the SuperDARN data.

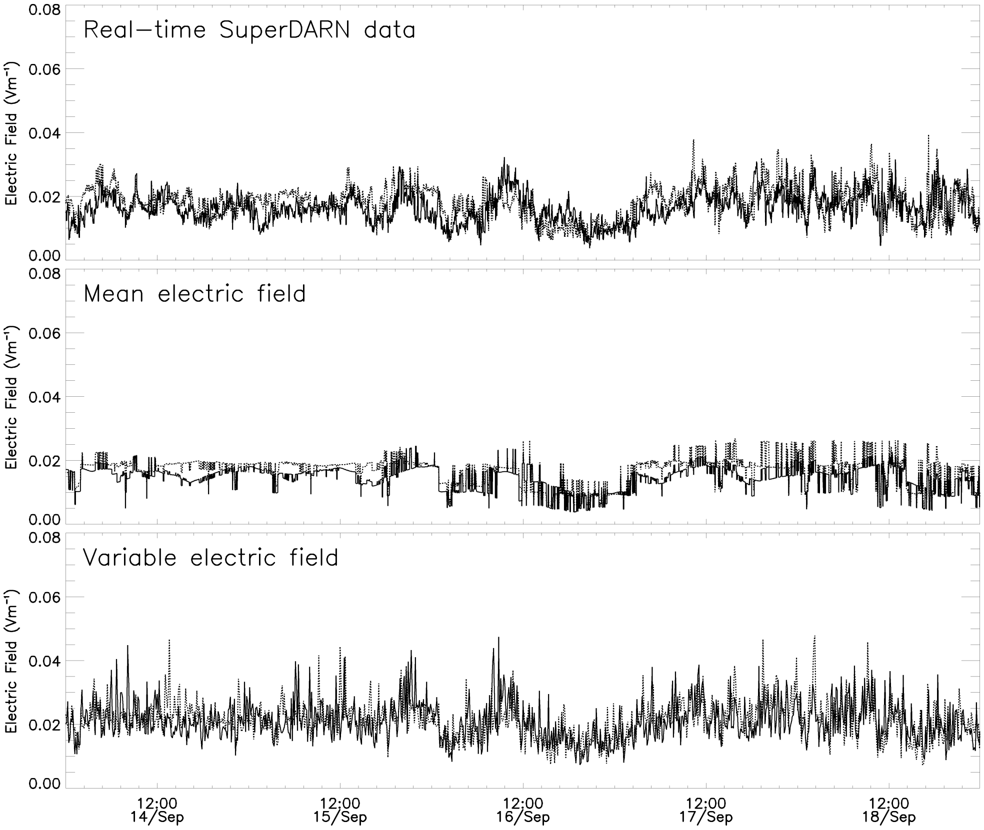
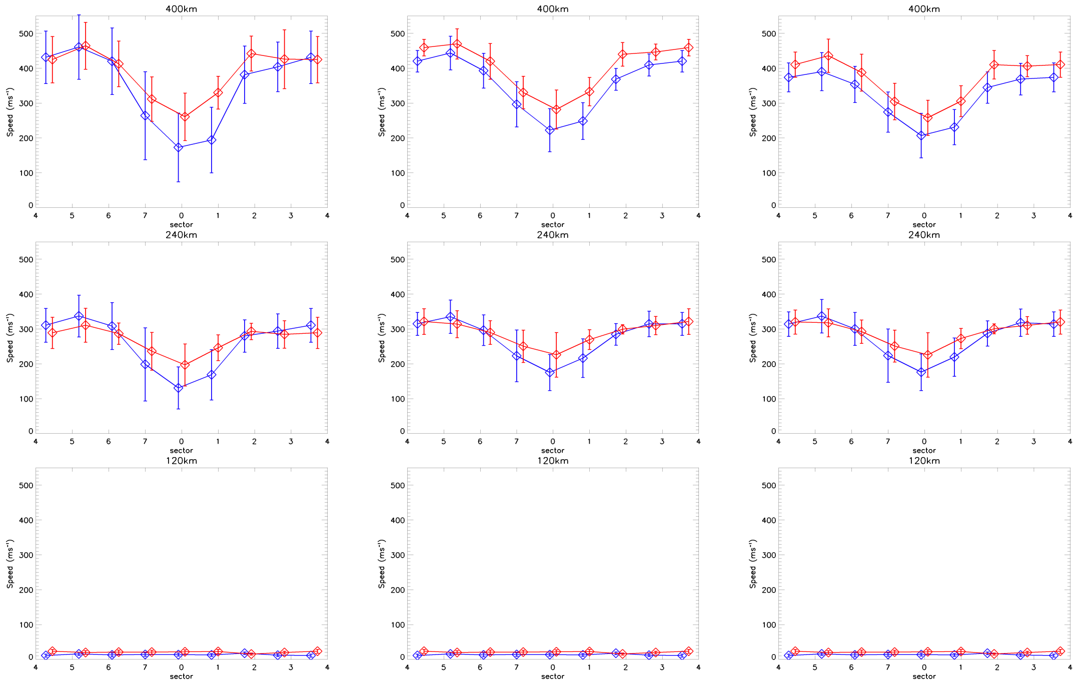


Figure 8.4: The same format as figure 8.3, but for the September 2002 simulation.

Figure 8.4 is of the same format as figure 8.3, but for second simulation during $14-18^{th}$ September 2002. The SuperDARN observations in both hemispheres show an electric field strength of around 0.02Vm^{-1} consistently, with a slight drop 1200UT on 16^{th} and 17^{th} September. The features identified in the previous simulation are also seen here. Firstly, the general trend in the mean and variable model follows that of SuperDARN, however the variable model peaks at greater values. Secondly, the mean model shows very static behaviour with step changes corresponding to IMF changes. Thirdly, the variable model shows a more similar time variation to the SuperDARN data, however the frequency is much greater.

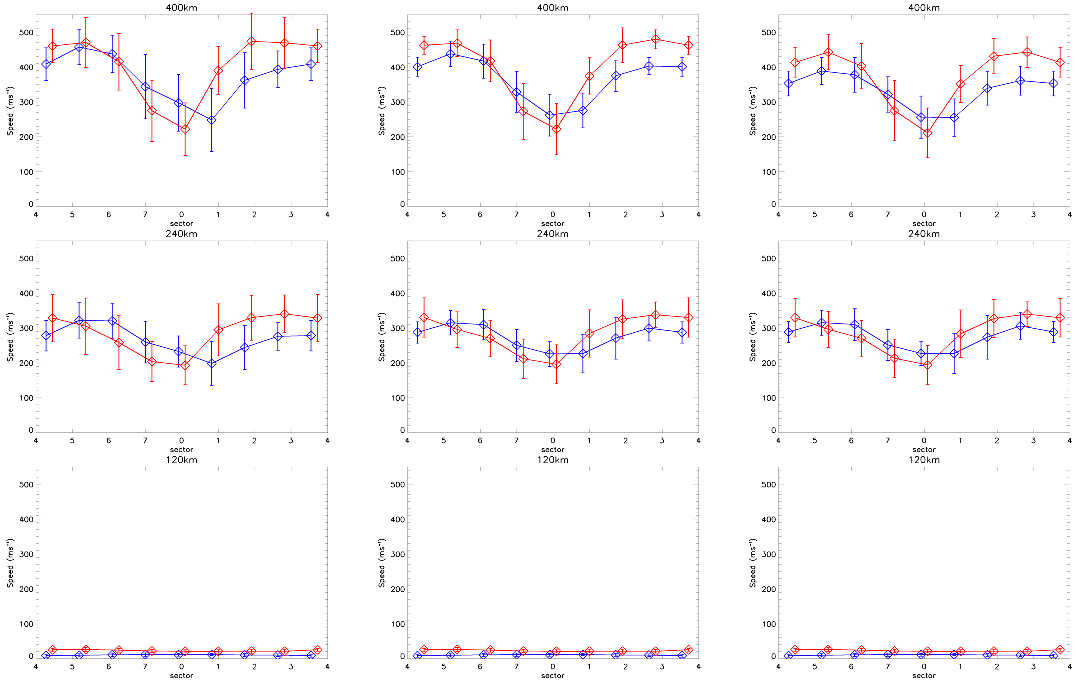
8.2.2 Neutral Wind Speeds



(a) SuperDARN (b) Mean (c) Variable

Figure 8.5: CMAT2 neutral winds from the March 2001 simulation, binned as a function of IMF sector. The wind speeds are averaged within the region $|\phi_m| > 80^\circ$ in each hemisphere at 120km, 240km and 400km. Figure 8.5(a) is from CMAT2 driven directly by SuperDARN and figures 8.5(b) and 8.5(c) are from CMAT2 using the IMF-dependent mean and variable electric field models. In each case the NH is shown in blue and the SH in red.

The mean neutral winds from the March 2001 simulations, as a function of IMF sector and altitude, are shown in figure 8.5. The left-hand plots (a) show results from CMAT2 driven by SuperDARN electric fields, whilst the middle (b) and right-hand (c) plots use the mean and variable models, respectively. The SuperDARN simulation shows a strong increase in winds when $B_z < 0$ and the overall speeds are slightly greater in the SH. There is also a small dependence on B_y, particularly at 240km, which caused an increase in SH speeds when $B_y > 0$ and a decrease when $B_y < 0$. The B_z dependence of both the mean and variable models is consistent with the SuperDARN model, as are the absolute values in each sector. This is with the exception of the NH speeds in sector 0, which are about 50kms^{-1} greater in the mean and variable models. The B_y dependence on SH speeds is seen less strongly in both the mean and variable models, where little hemispheric asymmetry is apparent.



(a) SuperDARN (b) Mean (c) Variable

Figure 8.6: The same format as figure 8.5, but for the September 2002 simulation.

Figure 8.6 is of the same format as figure 8.5, but shows results from the September 2002 simulation. The SuperDARN-driven version of CMAT2 shows the same B_z dependence as the previous simulation interval, where negative B_z produces greater speeds. However, the SH speeds are no longer seen to be greater in sector 0. The B_y dependence on the NH-SH differences is much more apparent in this simulation, particularly at 240km. Again, these general features are found in both the mean and variable models; the B_z dependence is quite clear and the absolute values of neutral wind speeds are consistent between all three models. The B_y dependence is seen in both the mean and variable models, particularly at 240km.

8.2.3 Joule Heating

Figures 8.7 and 8.8 show the total hemispheric, and global, Joule heating from the March 2001 and September 2002 periods respectively. During both simulation periods similar behavior is seen in each of the SuperDARN-driven simulations, where negative values of B_z correspond to strongly increased heating, peaking at around 120GW globally. There is a hemispheric dependence of heating on the sign B_y, however this is opposite to that seen in the winds. A negative B_y value results in stronger heating in the SH and a positive

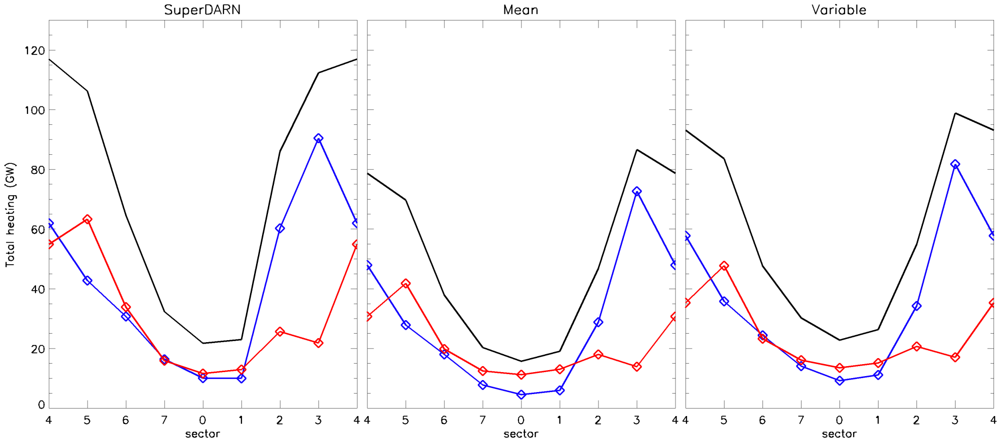


Figure 8.7: Joule heating averaged as a function of IMF sector from CMAT2 driven by the three different electric field models, during the March 2001 simulation. The left hand plot represents the SuperDARN driven run and the middle and right hand plots show results from the mean and variable electric field models. Blue and red represent the average of the total hemispheric power in the NH and SH, respectively. The black line is the total global power.

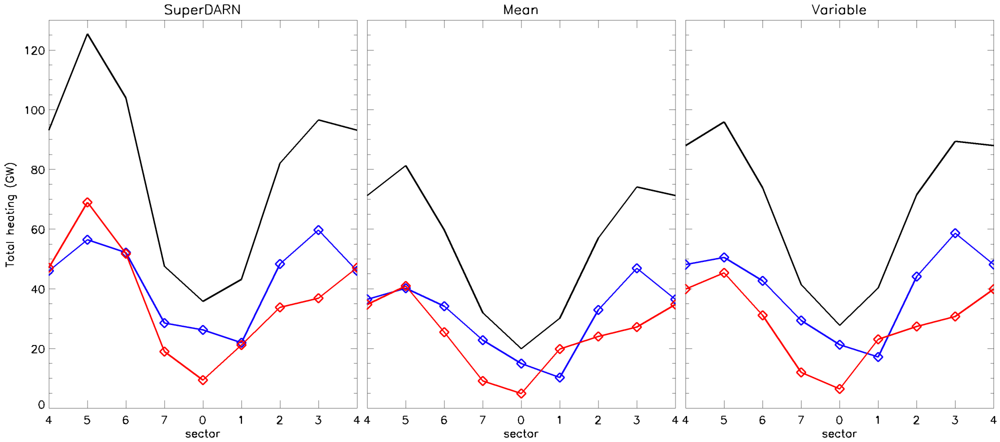


Figure 8.8: The same format as figure 8.6, but for the September 2002 simulation.

value produces stronger heating in the NH. In each of the simulations driven by the mean and variable electric field models, both the B_y and B_z dependence is consistent with these findings. However, the magnitude in heating is lower when using either of these models to drive CMAT2, than when using SuperDARN. This difference is smaller using the variable model, which does succeed in producing more heating than using just the mean electric field. This difference between all three model is seen most clearly in sectors where B_z is negative.

8.2.4 IMF Magnitude

	IMF sector							
Period	0	1	2	3	4	5	6	7
16–20 March 2001	5.5	4.3	10.3	11.5	11.6	9.3	7.2	6.8
14–18 September 2002	9.2	6.3	6.4	4.7	6.3	10.1	10.0	9.5

Table 8.1: Mean IMF magnitude (nT) during the two five-day simulation intervals, as a function of IMF sector.

	IMF sector							
E-field model	0	1	2	3	4	5	6	7
mean (2001)	73%	83%	54%	77%	67%	66%	59%	63%
variable (2001)	105%	115%	64%	88%	80%	79%	74%	94%
mean (2002)	56%	70%	70%	77%	77%	65%	57%	67%
variable (2002)	78%	93%	87%	93%	94%	76%	71%	87%

Table 8.2: Relative heating as a function of IMF sector with respect to the SuperDARN-driven simulations for the same period. The values represent total global power, which is shown by the black lines in figures 8.7 and 8.8.

The mean values of IMF magnitude are calculated as a function of sector during both five-day simulation intervals and tabulated in table 8.1. During the 2001 simulation, the mean IMF magnitude exceeds 7.5nT (the highest threshold for the IMF-driven models) in sectors 2–5, when B_z is strongly negative. During the 2002 IMF values in excess of 7.5nT are seen for sectors 0 and 5–7, when B_y is negative. Table 8.2 shows tabulated values from figures 8.7 and 8.8, which is the amount of heating from the mean and variable simulations, as a function of IMF sector, relative to the SuperDARN simulations. The variable model overestimates the heating in sectors 0 and 1 during the 2001 simulations, where the very lowest heating occurs. For all other elements of table 8.2 the IMF-driven models produce a lower estimate of heating than that those driven using SuperDARN. The sectors in table 8.1, which show the strongest IMF values correspond to the sectors in table 8.2 where the Joule heating is underestimated most strongly when using the mean and variable electric field models. The significance of this is that the SuperDARN data used to drive CMAT2 directly are based on observations under very strong IMF conditions. Conversely, the mean and variable electric field models are based on only three levels of IMF strength. Therefore any IMF input above 7.5nT will produce the same electric fields, whilst in the SuperDARN-driven model values much greater than 7.5nT are causing much greater heating.

	IMF sector							
E-field model	0	1	2	3	4	5	6	7
mean NH (2001)	128%	128%	97%	101%	97%	96%	94%	112%
mean SH (2001)	108%	101%	100%	105%	108%	101%	102%	106%
variable NH (2001)	120%	120%	90%	91%	87%	85%	84%	104%
variable SH (2001)	99%	93%	93%	95%	97%	94%	94%	98%
mean NH (2002)	88%	111%	104%	102%	98%	96%	95%	95%
mean SH (2002)	100%	96%	98%	102%	100%	100%	101%	99%
variable NH (2002)	86%	103%	94%	92%	86%	85%	86%	93%
variable SH (2002)	95%	90%	91%	94%	90%	94%	97%	100%

Table 8.3: Relative neutral wind speeds ($|\phi_m| > 80^\circ$) as a function of IMF sector with respect to the SuperDARN-driven simulations for the same period. The values are compared at an altitude of 400km, which corresponds to the top row of plots in figures 8.5 and 8.6.

Table 8.3 displays the relative hemispheric mean speeds (from $|\phi_m| > 80^\circ$ in each hemisphere) from each of the mean and variable simulations as a function of IMF sector. These values are expressed as a percentage, relative to the values from the corresponding SuperDARN-driven simulation. The values are consistently close to 100%, suggesting that both the mean and variable models agrees well with the SuperDARN model. In the NH, the sectors showing the greatest discrepancy with the SuperDARN simulations are the sectors where the IMF magnitude, in table 8.1 are greater.

8.2.5 Discussion

An example of the electrostatic potential from SuperDARN over a ten minute period is compared to the two IMF-driven models in figure 8.2. It was observed that the IMF parameters remain relatively static during this period, which causes the mean electric field model to remain at the same value. Instead, the observed SuperDARN values are seen to change somewhat during this period. The variable electric field model attempts to reproduce this feature, however, the variability appears to be too rapid because there is no correlation between successive time-steps. This causes the model to vary too quickly, when compared to SuperDARN, despite the two possessing the same covariance properties. In the time series in figures 8.3 and 8.4 this effect is also seen, where the SuperDARN field appears to change smoothly, whilst the variable model is more erratic. The mean electric field model does, of course, also vary but this only occurs during significant IMF changes and appears as large steps in electric field strength. Ridley et al. (1998) showed that changes in ion convection, as a result of changing IMF, could take approximately 10–20 minutes and that the transition was usually smooth. The variable model therefore needs to account for this by quantifying the temporal, as well as spatial, variability in order to produce more realistic electric fields.

In figures 8.5 and 8.6, the results of each of the three electric field models on CMAT2 neutral wind speeds are shown. Between both the SuperDARN driven runs, three main features were identified: firstly, greater speeds in the SH compared to the NH, secondly, an increase winds when $B_z < 0$ and thirdly, a hemispheric dependence on the sign of B_y. Of these features only the first two are seen clearly in the March 2001 simulation and only these last two are seen clearly in the September 2002 simulation. In each of the mean and variable electric field simulations, these same features can be identified. The slightly greater speeds in the SH, which agree with observations by Thayer and Killeen (1993), are thought to result from the greater values of both $|\mathbf{E}|$, in SuperDARN's SH, and $|\mathbf{B}|$, in CMAT2's SH, via $\mathbf{E} \times \mathbf{B}$ drift. The B_z dependence is identified as an obvious consequence of the increased field strength under these conditions, as is seen in the Weimer (2005) model. Lastly the hemispheric dependence on B_y results from the rotation of the convection pattern as a function of the sign of B_y. As discussed in the previous chapter, this is caused by the interaction between the pressure gradient and ion convection, as observed by Förster et al. (2008). It should therefore be expected that both the mean and variable electric field models reproduce these three features, because they are designed with an IMF dependence.

In the previous chapter it was concluded that the electric field variability should have little effect on wind speeds, which is supported here by the similarity between the speeds seen in the mean and variable electric field driven simulations. In both the 2001 and 2002 periods, the two models produce nearly identical winds at 240km, whilst the variable model produces lower values at 400km, by around 10%. This is likely to be a result of the rapidly changing direction of the convection pattern seen in the variable model. The effect of this is that there is insufficient time for the neutral atmosphere to accelerate in the direction of the ion convection, before this direction changes. As a result, the neutral atmosphere is unable to gain as much momentum when driven by the variable model. It has been identified that whilst time variation in electric fields is observed using SuperDARN, it is in fact on a slower scale than that of the randomly-varying model. This may explain the greater similarity between the SuperDARN-driven model and that driven by the mean electric fields, when simulating neutral wind magnitudes.

Lastly, figures 8.7 and 8.8 compared the hemispheric and global Joule heating observed in the March 2001 and September 2002 simulations, respectively. In this case two features were identified in the SuperDARN driven runs. Firstly, a large increase in heating when $B_z < 0$ and secondly, a dependence on the sign of B_y, which was opposite in each hemisphere. The B_z dependence is again explained by the increase in the value of $\mathbf{E}$ measured by SuperDARN under these conditions. CMAT2 therefore produces more heating through $Q_J = \Sigma_P E^2$ from (3.7). The B_y dependence of heating in each hemisphere was seen to increase in sectors where the SuperDARN electric fields were most variable. The SuperDARN, mean and variable simulations all differ on the magnitude of Joule heating. As expected, the lowest values are seen in the model using the mean electric fields, which contain the least variability. However, this model is still driven using solar wind input every two-minutes and so is far more variable than that of the Foster model. The resulting underestimate of Joule heating is therefore not as significant. Whilst the variable electric field model is seen to produce greater Joule heating than the mean model, it is still slightly less than is seen in the SuperDARN-driven simulations. One cause of this may be that the random variables used to drive this model are taken from a normal distribution. However, it was shown in chapter 7 that the SuperDARN cross-cap potential distributions are consistently non-Gaussian. The type of distribution varies depending on IMF conditions, but are generally more sharply-peaked and often heavy-tailed. It is likely that using Gaussian random variables to drive these fields will in fact result in incorrect heating within the model.

Figure 8.1 shows the IMF magnitude and sector for both five day simulation intervals. The IMF-driven models are driven by three different strength levels: $|\mathbf{B}| < 5.0$nT, $5.0\text{nT} \leq |\mathbf{B}| < 7.5$nT and $7.5\text{nT} \leq |\mathbf{B}|$, whilst the true IMF can sometimes exceed these values significantly. Therefore the mean and variable models under an IMF strength of, for example, 10nT and 20nT would produce the same electric field. It can be seen from figure 8.1 that, during the 2001 simulation interval, the IMF strength can be seen to exceed 25nT. The SuperDARN driven run, under these extreme conditions, is therefore expected to produce significantly more heating. Table 8.1 shows the mean IMF magnitude that occurred in each sector during the two simulations and table 8.2 shows the amount by which the IMF-driven models underestimate total heating in each sector when compared to SuperDARN-driven simulations. It is found that there is a correlation between these two effects, supporting the explanation of this discrepancy.

Although the variable model shows an improvement over simply using a mean electric field model with regards to Joule heating, it is still an oversimplification of much more complicated system. The fact that it does not contain any measure of temporal variability means that it varies too rapidly and the basis on three levels of IMF strength mean it is unable to simulate the response to the most extreme IMF conditions.

Summary

A new empirical electric field model has been developed, based on SuperDARN observations, which quantifies both the mean field and its variability as a function of IMF orientation and magnitude. The importance of this variability to model neutral wind speeds and Joule heating correctly in GCMs, particularly CMAT2, has been demonstrated in chapters 6 and 7. Three versions of CMAT2 have been run over two five-day intervals using different electric field inputs. One simulation uses real-time SuperDARN data from the period in question, whilst the other two are driven by the IMF values from the same period. One of these models is uses the mean electric field as a function of IMF strength and orientation, whilst the other also contains a measure of its spatial variability.

It is found that both the IMF-driven models were able to match the SuperDARN-driven model when reproducing neutral winds. When binned by IMF sector, mean wind speeds in the polar region were typically found to be within 10% of the SuperDARN-driven values at 400km altitude. Below this the values were very much closer. The variable model was found to underestimate wind speeds slightly, which is due to the rapidly changing convection patterns produced by the randomly generated variability. It is thought that these conditions do not allow the neutral atmosphere to gain as much momentum as in the other two models because they are not accelerated in the same direction for long enough. However, the difference between speeds produced by the mean and variable electric field models is little. The behaviour of all three models showed the same features as a function of B_y and B_z in each hemisphere.

When total hemispheric power from Joule heating was binned by IMF sector, similar behaviour was observed in the models. Again the dependence on B_y and B_z from each model appeared to show the same trends and again the IMF-driven models underestimated the heating produced by the SuperDARN model. It is thought that this was caused by the unusually large IMF magnitudes seen during the simulation intervals, particularly in the 2001 simulation. The largest IMF values observed fell in sectors where the relative heating from IMF-driven models was least compared to SuperDARN. The mean electric field model showed the lowest heating, which is thought to result from the reduced variability it possesses. The variable electric field model underestimated the heating somewhat, however it was able to produce more similar Joule heating values to the SuperDARN-driven model.

The variable electric field model demonstrated in this chapter has had some success in producing increased Joule heating, however it possesses several weaknesses. The first is that it is may only be driven by three different levels of IMF strength, which can result in an underestimate of the electric field during the most extreme IMF conditions. An obvious solution of which would be to divide the data into smaller bins and increase the number of levels, however the limited quantity of SuperDARN data, particularly for the strongest IMF conditions, makes this difficult. Secondly, the variability quantified in the covariance matrix contains a measure of how each convection pattern varies spatially but none of how they evolve with time. It is thought that this causes the electric field to vary too rapidly, which has an adverse effect on neutral wind speeds. Thirdly, the model relies on Gaussian random variables, which are not a true representation of the variability in SuperDARN. Finally, the underestimation of the magnitude of electrostatic potential from SuperDARN observations will also be present in this new model and will result in further underestimates of heating and wind speeds.

This page was intentionally left blank

Chapter 9

Conclusions

9.1 Summary of Findings

9.1.1 Estimating CME Densities using STEREO HI

Chapters 2 and 3 present a method for estimating solar wind and CME densities using the STEREO HI observations. Density distributions are calculated in the ecliptic plane via inversion of HI intensities based on Thomson scattering theory. For each pair of HI images, from STEREO-A and -B, a density map is calculated in the region of the ecliptic common to fields of view of both instruments. By applying this technique to successive pairs of images, the time evolution of density within this region may be determined. The technique is first applied to the quiet solar wind and average densities are found to be consistent, as a function of heliocentric distance, with in situ measurements from the two HELIOS spacecraft. The technique is then applied to two Earth directed CMEs and estimates of their density and speed are made. These estimates are extrapolated to a distance of 1AU, based on established CME evolution. In both cases the velocity estimates are found to be consistent with in situ measurements by the WIND spacecraft at L1, however the density estimates are less accurate. This type of tomographic inversion of white-light images has been applied to the corona and CMEs in many previous studies, particularly since the launch of the STEREO mission. However, it is typically performed on data from coronagraphs, which observe a region much closer to the Sun than HI. This is because the electron-scattered K-corona has a greater relative intensity compared to the dust-scattered F-corona in this region. Further from the Sun, in the HI field of view, the F-corona begins to dominate by a factor of approximately 100. The tomographic inversion of HI data therefore requires the K-corona to separated from the F-corona and background star-field. Due to uncertainties in this subtraction, a correction must be applied to the data to produce an inverse problem, which can be solved to provide physically consistent densities.

The method suffers several limitations based on the orientation of the spacecraft and objects observed in the HIs. The passage of planets, comets and the galactic plane through the field of view of either HI instrument makes K-corona measurements impractical. As a result the method may not be applied to periods of HI data when this is the case. As the positions of the two spacecraft evolve throughout the mission, so too does the angle at which the HI lines of sight intersect. A separation angle of 90° is ideal for tomography because it becomes easier to resolve line of sight ambiguities. However, when the separation approaches 180° these ambiguities become significant. With the absence of polarised brightness measurements, two spacecraft is the fewest required for the tomographic inversion to be performed. However the two points of view is insufficient to determine detailed density structures, such as CMEs, and so they appear somewhat amorphous in the reconstruction. The method may only be performed in the region common to the field of view of both spacecraft, which limits the coverage available from just two spacecraft. A combination of the factors listed above means that the method has only been applied to two CMEs throughout the entire STEREO mission.

9.1.2 Modelling the Thermospheric Effects of Electric Field Variability CMAT2

Chapter 5 presents a modelling study of the effects of ionospheric electric field variability on thermospheric energy and momentum transfer. Two models of electrostatic potential are considered; the empirical Foster model and a model based on coherent scatter radar measurements from SuperDARN. It was shown that the K_p index, used as a proxy to drive the Foster model, was poor at reproducing the same cross-cap potential as is seen in SuperDARN. For the highest K_p values Foster produced a significantly higher cross-cap potential than SuperDARN. This is be cause the K_p index represents a three-hour maximum in magnetometer readings, which means it is unable to reproduce any changes on a smaller time-scale, like those seen in SuperDARN. As a function of IMF sector, the

the Foster model showed almost no change in cross-cap potential, whilst that seen in SuperDARN is strongly influenced by the B_z component. The results were also compared to the model by Weimer (2005), for average solar wind conditions. SuperDARN was found to match the Weimer model well for most IMF conditions but was significantly lower when $B_z \ll 0$. Cousins and Shepherd (2010) found the same effect when comparing SuperDARN to other models and observations, which Gillies et al. (2011) attribute to incorrect assumptions in the refractive index of back-scatter regions used in determining velocities from SuperDARN.

Two CMAT2 simulations are performed over the same five-day period, 16^{th}–20^{th} March 2001. One model is driven using Foster's model, determined from the K_p index, and the second uses SuperDARN electric fields. The resulting Joule heating is found to follow the relationship $Q_J \propto e_m^2 + \sigma^2$ proposed by Codrescu et al. (2000), with Foster's model producing lower Joule heating in the model than SuperDARN. The neutral winds, however, are found to be independent of variability and are proportional to the mean electric field strength. In order to accurately simulate energy and momentum transfer to the neutral atmosphere, it is therefore necessary to use an electric field model that contains temporal variability and is based on a proxy, or proxies, that accurately predict the strength of the electric field.

9.1.3 Hemispheric Asymmetry in Ionosphere-Thermosphere Coupling

Chapter 6 studies the asymmetry in energy and momentum transfer between the ionosphere and thermosphere in CMAT2. The model is run with SuperDARN electric fields for two five-day periods, $16^{th} - 20^{th}$ March 2001 and $14^{th} - 18^{th}$ September 2002, during equinoxes at solar maximum. The neutral winds in the southern hemisphere were found to be slightly greater than the north. This effect is due to a stronger magnetic field (from the IGRF) and a greater electric field strength (from SuperDARN) producing a greater $\mathbf{E} \times \mathbf{B}$ drift in the south. The neutral wind speeds increase strongly when B_z is negative compared to when it is positive, which is consistent with established behaviour. There is also a dependence on the B_y component. The speeds are lower in the NH and SH when B_y is positive or negative respectively. This is thought to be caused by the clockwise rotation of the convection pattern seen under these IMF conditions. This acts to oppose the pressure gradient, reducing the net acceleration on the neutral atmosphere. This effect is seen in observations from CHAMP, which are binned in the same way for each hemisphere (Förster et al., 2008; Förster et al., 2012).

A study of Joule heating is performed by running two CMAT2 simulation over the five-days $16^{th} - 18^{th}$ March, one using SuperDARN electric fields and the other using Foster's model. It is again found that Joule heating has a strong dependence on electric field variability. When binned according to IMF sector, the greatest heating in SuperDARN driven simulations is observed in the sectors where the SuperDARN electric fields are most variable. The Foster fields, however produce far lower Joule heating, which results from their lack of variability. When compared with IS radar observations the magnitude of heating in the SuperDARN simulations is found to be similar as a function of IMF sector. However, the distribution of heating, in MLT and magnetic latitude, in the Foster driven simulations is found to match the radars observations better. This is because both the Foster and SuperDARN electric fields are using the same particle precipitation model, which was designed to work with Foster's (K_p dependent) model. The overall values in neutral speeds and heating are found to be lower when compared to other studies by Förster et al. (2008) and Knipp et al. (2005) respectively. Cousins and Shepherd (2010) found SuperDARN can underestimate the cross-cap potential by 71% (NH) and 63% (SH) when compared to other observations, which may explain the low neutral winds and Joule heating seen in CMAT2.

9.1.4 IMF-Driven CMAT2 Simulations

In chapter 7 a new empirical electric field model has been developed, based on SuperDARN observations, which quantifies both the mean field and its variability as a function of IMF orientation and magnitude. The purpose of this is to avoid the lack of spatial variability present in averaged electric field models, such as those of Foster and Weimer, which are known to contribute to low Joule heating in GCM studies. This new model bins eleven years of SuperDARN observations, from 2000 to 2010, as a function of IMF strength and orientation. The mean electrostatic potential distribution is determined over the CMAT2 grid for each IMF bin, as is the covariance matrix of the individual potential values over the grid. The significance of this is that the covariance matrix contains a measure of the spatial variability of the SuperDARN electric fields. As a result it may be used to produce a randomly varying electric field within the CMAT2 that possesses the same covariance properties as the electric fields observed by SuperDARN. Two versions of this model are used to drive CMAT2; the first is the variable model just described and the second suppresses the variability and uses only the mean electric field. The purpose of this is to directly study the difference that arises from this variability. CMAT2 is driven over the same two five-day periods used previously; $16^{th}-20^{th}$ March 2001 and $14^{th}-18^{th}$ September 2002. The neutral wind speeds from both these models are found to be consistent with the simulations driven directly by SuperDARN, although lower than speeds observed by CHAMP at the same altitude. An increase in wind speeds with negative B_z is observed, as is the hemispheric dependence on B_y identified in the previous chapter. When Joule heating is compared, both the mean and variable electric field models are found to underestimate heating compared to the model driven directly by SuperDARN. The variable model does produce more heating than the model which only uses mean electric fields. The strong increase in Joule heating with negative B_z that was established in the previous chapter is seen in both models, as is the hemispherically asymmetric dependence on B_y. The variable model is binned according to three levels of IMF strength, the greatest being $|\mathbf{B}| > 7.5$nT, and so is unable to reproduce the high Joule heating seen in the SuperDARN simulation when the $|\mathbf{B}|$ peaks at over 20nT. The variable model contains no measure of temporal variability and the electric field values vary much more rapidly than is seen in real SuperDARN data, which has an adverse effect on neutral wind speeds.

9.2 Improvements to Models and Future Studies

9.2.1 Estimating CME Densities using STEREO HI

Given that the method is limited by the restricted coverage of only two spacecraft, the most obvious improvement would be to increase the number of spacecraft available. Whilst this may be an unrealistic aspiration, a number of new solar missions are expected to launch in the coming years. Both Solar Probe Plus and Solar Orbiter are scheduled for launch in 2018 and posses heliospheric imaging capabilities. Combined with the existing STEREO-A instrumentation, improved coverage of the inner heliosphere may one day allow more accurate tomography methods to be performed.

Without the addition of more spacecraft there are still many ways in which the techniques presented here may be developed. The existing STEREO data contain 19 Earth-directed CMEs (table 5.2), two of which have been studied in this thesis. The remaining 17 were excluded on three bases; spacecraft orientation, obstructing planets and the presence of the galactic plane in the background. Events 8-15, occurring in 2010-11, were rejected because the STEREO spacecraft were separated such that the HI-1 fields of view were at approximately 180° to each other. This introduces line of sight ambiguities in the tomographic inversion, but could potentially be addressed by the inclusion of the third point of view provided by SMEI on-board Coriolis in Earth orbit. During the years 2010-11 these three view points would have been separated by approximately 90°. Whilst 90° is not ideal for tomography, three-spacecraft at this separation provide far better coverage than two. Further to this, a third point of view would increase the region of space in which the technique can be applied, increasing the number of CMEs available. Secondly, events 17-24 were rejected on the basis of planets in one or more HI-1 fields of view. The study performed here has been restricted to two-dimensions for the purpose of simplicity, to demonstrate the method. With Earth-impacting events of interest, the ecliptic plane was chosen because it contains both spacecraft and the Earth. A consequence of this is that planets frequently appear in this part of the image. Extending the technique to three-dimensions would largely reduce this limitation. Lastly, the presence of the galactic plane in the field of view caused events 6,7 and 22 to be rejected. Due to its large intensity, this is a more difficult problem to address.

A great deal of work has focused on tomography of the corona near the Sun's surface. The biggest obstacle in performing the technique at large elongation angles lies in identifying the contribution of the K-corona to the total observed intensity. The main improvement to this technique would therefore be a more thorough method for separating the relative contributions from each source.

Given the limitations of having two view points, a significant improvement to the technique would come from combining it with forward-modelling, like that of Frazin et al. (2009). This requires assumptions and measurements of CME morphology, which are studied in detail by Wood et al. (2009); Tappin and Howard (2009c) and Antunes et al. (2009); Davies et al. (2012); Davies et al. (2013), respectively.

9.2.2 Modelling the Thermospheric Effects of Electric Field Variability and Hemispheric Asymmetry in CMAT2

A weakness of the SuperDARN driven simulations was that they produced a lower value of both neutral wind speeds and Joule heating when compared to observations and other studies. This is expected to be a result of SuperDARN underestimating the electric field magnitude. If the source of this underestimation could be quantified and corrected for then the model should be capable of producing a neutral atmosphere more consistent with observations. The strong Joule heating in the NH was found to result from high variability in the SuperDARN electric fields. This variability was seen less in the SH, due to its greater dependence on the empirical model when observations are lacking. Greater radar coverage in the SH would therefore reduce this problem, providing similar measurements to the NH. Likewise, the lack of IS radar measurements in the SH meant a comparison of CMAT2 to observations was only possible in the NH. To properly investigate the hemispheric differences in Joule heating, this comparison is also necessary in the SH. The magnitude of heating in the SuperDARN driven simulations was found to be close to that seen in the radars, but its spatial distribution was not. In order to produce more realistic Joule heating in CMAT2, the SuperDARN electric fields require a matching particle precipitation pattern.

9.2.3 IMF-Driven CMAT2 Simulations

The variable electric field model is based on SuperDARN observations and so also suffers from the same problem of underestimated potentials. A means of correcting for this would again be a necessary improvement the model. Again, the same particle precipitation pattern has been used in conjunction with this electric field model due to a lack of better alternatives. It would be necessary to include an improved particle model in CMAT2, in order to correctly simulate Joule heating. A means to reduce the rapid fluctuation of the variable model would be expected to produce electric fields that show a more similar variation to SuperDARN. A simple method to achieve this may be to read electric fields in to CMAT2 less frequently and to interpolate between them, to emulate the smooth evolution of the SuperDARN fields. A more rigorous approach would be to quantify the temporal variability of the electrostatic potential, as well as its spatial variability, such that the model could create a random pattern but with a realistic, smooth time variation.

Summary

Currently, space weather prediction relies on observations from within a few $R_\odot$, typically coronagraphs, which are then propagated outward from the Sun using MHD modelling. The solar wind will usually take several days to travel from the Sun to 1AU and therefore properties such as speed and density are likely to evolve significantly over this distance. The work presented here demonstrates how these remote sensing observations may be extended to much larger distances from the Sun, using heliospheric imagers, and therefore provide a potential means to constrain forecasts as CMEs propagate towards Earth. Possible future missions with greater coverage and improved instruments, combined with more thorough image processing methods, would reduce many of the limitations that were identified.

Space weather forecasting of terrestrial effects is typically performed by using proxies, such as solar wind speed, density and magnetic field, to drive electric fields in a GCM, which can simulate energy and momentum transfer between the ionosphere and neutral atmosphere. The importance of electric field variability in controlling this energy transfer has been demonstrated, in addition to the hemispheric dependences that exist. Based on this an empirical IMF-dependent model has been developed to drive CMAT2, which has been shown to reduce the amount that Joule heating is underestimated, a common problem in GCMs. Although this technique also has its limitations, it demonstrates a means to reduce the heating underestimation by quantifying the spatial variability of the observed ionospheric electric field.

Successful space weather forecasts depend both on accurate predictions of Earth impacting solar wind properties and on improved modelling of how the thermosphere responds to them. Improved techniques for estimating these properties from remote sensing are therefore a necessity, as are improved models of how they drive energy and momentum transfer within the Earth's atmosphere.

Appendices

Appendix A

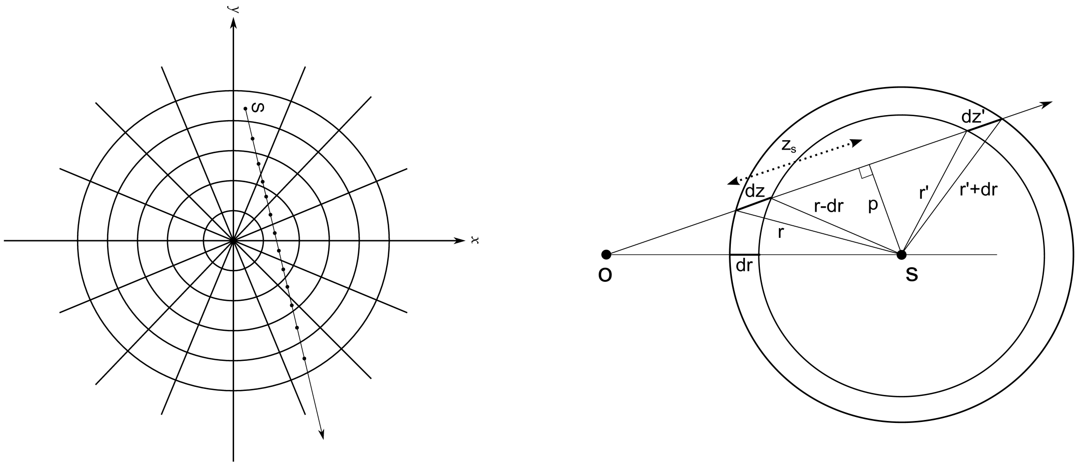
Heliospheric Grid

The determination of the observation matrix **H** in chapter 4 requires the path of each HI line-of-sight to be determined through the heliosphere. This is performed using the method of Tappin and Howard (2009b), which is performed via an iterative process over a three-dimensional heliospheric grid. At each step in the iteration the distance to the next cell in each of the three coordinate dimensions is computed and the shortest of these is determined to be the next cell that the line-of-sight crosses.

In this thesis a three-dimensional heliocentric grid measuring $0.025AU$ radially, 3° longitudinally and 3° azimuthally has been used. These coordinates are denoted r, ϕ and ψ, respectively. The grid is aligned to a Heliocentric Earth Equatorial (HEE) coordinate system, such that ϕ lies in the $x - y$ plane. For a given pixel, the spacecraft position in the HEE grid is determined and the direction of the pixel's line-of-sight from that position is calculated. The method of determine the distances in each coordinate dimension then follows that of Tappin and Howard (2009b).

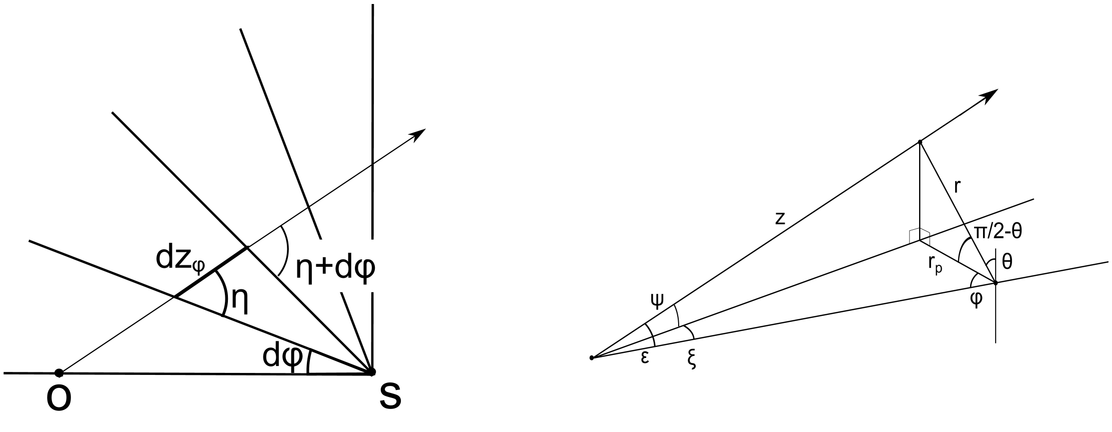
Figure A.1 shows the method applied to determining the distance in each dimension. The radial step (figure A.1(b)) is calculated from the following, depending on whether the line-of-sight is approaching or moving away from the Sun, where p is the closest distance between the line-of-sight and the Sun,

$$dz_r = \begin{cases} (r+dr)^2 - p^2]^{1/2} - z_s & \text{if } z_s < 0 \\ (r+dr)^2 - p^2]^{1/2} - z_s & \text{if } z_s > 0 \\ 2(r^2 - p^2)^{1/2} & \text{if } p > r - dr \end{cases} \tag{A.1}$$



(a) Line-of-sight through the heliocentric grid.

(b) Determination of the distance to the next radial boundary.



(c) Determination of the distance to the next longitudinal boundary.

(d) Determination of the distance to the next azimuthal boundary.

Figure A.1: The geometry used in determining the line of sight through the heliocentric grid. Reproduced from Tappin and Howard (2009b).

The azimuthal step (figure A.1(d)) is given by

$$dz_\phi = \frac{r_p \sin d\phi}{\sin(\eta + d\phi)} \frac{1}{\cos\psi} \tag{A.2}$$

Finally, the longitudinal step (figure A.1(c)) is calculated using $dz_\theta = z(\theta_1) - (\theta_2)$ from

$$z(\theta) = \frac{\cos\epsilon \pm [\cos^2\epsilon + (\sin^2\psi \tan^2\theta - \cos^2\psi)]^{1/2}}{\cos^2\psi - \sin^2\psi \tan^2\theta} \tag{A.3}$$

where the root is positive if $|\phi| > \pi/2$ and negative otherwise. Once the shortest of these distances is determined, the model moves this distance along the line-of-sight and the process is repeated using the new coordinates. The value of $\mathbf{H}$ is determined at the midpoint through each cell and the process is terminated after

$$z \geq 2(1 + \sin(\epsilon/2)) \tag{A.4}$$

Appendix B

LDLT Decomposition and Gaussian Random Variables

B.1 Matrix Decomposition

The covariance matrix, $\mathbf{\Sigma}$, is, by definition, a square, symmetric and positive definite matrix. It is therefore a relatively uncomplicated task to find the matrix, $\mathbf{A}$, which satisfies

$$\mathbf{A}\mathbf{A}^{\mathsf{T}} = \mathbf{\Sigma} \tag{B.1}$$

$\mathbf{A}$ may be found via a number of methods, the simplest of which is Cholesky decomposition. This involves decomposing $\mathbf{\Sigma}$ in to a lower and an upper triangular matrix, $\mathbf{L}\mathbf{U} = \mathbf{\Sigma}$, where $\mathbf{U} = \mathbf{L}^{\mathsf{T}}$, providing the solution by $\mathbf{A} = \mathbf{L}$. LDLT decomposition is a variant of this method, which is used instead because it is found to be more computationally stable. This involves decomposing $\mathbf{\Sigma}$ in to an upper and lower triangular matrix and a further diagonal matrix, $\mathbf{D}$, which satisfy

$$\mathbf{\Sigma} = \mathbf{L}\mathbf{D}\mathbf{L}^{\mathsf{T}} \tag{B.2}$$

$\mathbf{A}$ is then determined by

$$\mathbf{A} = \mathbf{L}\mathbf{D}^{1/2} \tag{B.3}$$

where $\mathbf{D}^{1/2}$ is simply a new diagonal matrix, each component of which is the square root of the corresponding component in $\mathbf{D}$. Because $\mathbf{\Sigma}$ is positive-definite, the diagonal elements of $\mathbf{D}$ are all positive. The elements of the $N \times N$ matrices $\mathbf{L}$ and $\mathbf{D}$ are determined by an iterative algorithm, which may be understood by considering their structure,

$$\mathbf{LDL}^{\mathsf{T}} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ l_{10} & 1 & 0 & \cdots & 0 \\ l_{20} & l_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{N0} & l_{N1} & l_{N2} & \cdots & 1 \end{pmatrix} \begin{pmatrix} d_0 & 0 & 0 & \cdots & 0 \\ 0 & d_1 & 0 & \cdots & 0 \\ 0 & 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & d_N \end{pmatrix} \begin{pmatrix} 1 & l_{10} & l_{20} & \cdots & l_{N0} \\ 0 & 1 & l_{21} & \cdots & l_{N1} \\ 0 & 0 & 1 & \cdots & l_{N2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \tag{B.4}$$

It is trivial to see that the first element of $\mathbf{D}$, d_0, is equal to the first element of $\mathbf{\Sigma}$, σ_{00}. The value of d_0 is then used to determine the first column of $\mathbf{L}$, where $l_{i0} = \sigma_{i0}/d_0$. The element, l_{10}, is then be used to find d_1, which is in turn used to solve for the elements in the second column of $\mathbf{L}$. Each successive element of $\mathbf{D}$ is determined by

$$d_i = \sigma_{ii} - \sum_{k=0}^{i-1} l_{ik} d_i \tag{B.5}$$

and elements of $\mathbf{L}$ are found using

$$l_{ij} = \begin{cases} \frac{1}{d_j}\left(\sigma_{ij} - \sum_{k=0}^{i-1} l_{ik} d_k l_{jk}\right) & \text{if } j \leq i \\ 0 & \text{otherwise} \end{cases} \tag{B.6}$$

This process continues until all the elements of $\mathbf{D}$ and columns of $\mathbf{L}$ have been determined. The result is used to calculate the matrix $\mathbf{A}$ for each IMF bin from (B.1).

B.2 Random Variables

The array of Gaussian random variables, $\mathbf{z}$, in equation (8.3) is necessary to create an electrostatic potential distribution from $\mathbf{A}$. These variables are produced using the Box Muller transform, which takes two independent random numbers from the uniform distribution between 0 and 1 and maps them to two independent Gaussian random numbers by

$$z_1 = \sqrt{-2\ln(x_1)}\cos(2\pi x_2) \tag{B.7}$$

$$z_2 = \sqrt{-2\ln(x_1)}\sin(2\pi x_2) \tag{B.8}$$

Two arrays, $\mathbf{x}_1$ and $\mathbf{x}_2$, of N uniform random numbers are used to produce two arrays of Gaussian random variables that are required in equation (8.3). To reduce computational expense within the model, it is convenient to calculate both $\mathbf{z}_1$ and $\mathbf{z}_2$ and apply them to opposite hemispheres.

This page was intentionally left blank

Bibliography

S. I. Akasofu. Several ‘controversial’ issues on substorms. *Space Science Reviews*, 113(1-2), 2004. ISSN 0038-6308. doi: 10.1023/B:SPAC.0000042938.57710.fb.

S.-I. Akasofu and Byung-Ho Ahn. Dependence of the amount of open magnetic flux on the direction of the interplanetary magnetic field. *Planetary and Space Science*, 28(5):545 – 547, 1980. ISSN 0032-0633. doi: http://dx.doi.org/10.1016/0032-0633(80)90035-5. URL `http://www.sciencedirect.com/science/article/pii/0032063380900355`.

Mark Allen, Jonathan I. Lunine, and Yuk L. Yung. The vertical distribution of ozone in the mesosphere and lower thermosphere. *Journal of Geophysical Research: Atmospheres*, 89(D3):4841–4872, 1984. ISSN 2156-2202. doi: 10.1029/JD089iD03p04841.

M. D. Altschuler. *Reconstruction of the global-scale three-dimensional solar corona*, page 105. 1979. doi: 10.1007/3-540-09417-2_4.

T. Amari, J. F. Luciani, J. J. Aly, Z. Mikic, and J. Linker. Coronal Mass Ejection: Initiation, Magnetic Helicity, and Flux Ropes. II. Turbulent Diffusion-driven Evolution. *ApJ*, 595:1231–1250, October 2003. doi: 10.1086/377444.

A. Antunes, A. Thernisien, and A. Yahil. Hybrid reconstruction to derive 3d heighttime evolution for coronal mass ejections. *Solar Physics*, 259(1-2):199–212, 2009. ISSN 0038-0938. doi: 10.1007/s11207-009-9409-7. URL `http://dx.doi.org/10.1007/s11207-009-9409-7`.

G. J. Bailey. The effect of a meridional E x B drift on the thermal plasma at L = 1.4. *Planetary Space Science*, 31:389–409, April 1983. doi: 10.1016/0032-0633(83)90154-X.

D. N. Baker, T. I. Pulkkinen, V. Angelopoulos, W. Baumjohann, and R. L. McPherron. Neutral line model of substorms: Past results and present view. *Journal of Geophysical Research: Space Physics*, 101(A6): 12975–13010, 1996. ISSN 2156-2202. doi: 10.1029/95JA03753. URL `http://dx.doi.org/10.1029/95JA03753`.

P. M. Banks and G. Kockarts. *Aeronomy.* 1973.

R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. Van der Vorst. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, 2nd Edition.* SIAM, Philadelphia, PA, 1994.

J. Bartels, N. H. Heck, and H. F. Johnston. The three-hour-range index measuring geomagnetic activity. *Terrestrial Magnetism and Atmospheric Electricity*, 44(4):411–454, 1939. ISSN 0096-8013. doi: 10.1029/TE044i004p00411. URL `http://dx.doi.org/10.1029/TE044i004p00411`.

D R Bates. The temperature of the upper atmosphere. *Proceedings of the Physical Society. Section B*, 64 (9):805, 1951.

S. Baumbach. Strahlung, ergiebigkeit und elektronendichte der sonnenkorona. *Astronomische Nachrichten*, 263(6):121–134, 1937. ISSN 1521-3994. doi: 10.1002/asna.19372630602. URL `http://dx.doi.org/10.1002/asna.19372630602`.

D. Bewsher, D. S. Brown, C. J. Eyles, B. J. Kellett, G. J. White, and B. Swinyard. Determination of the Photometric Calibration and Large-Scale Flatfield of the STEREO Heliospheric Imagers: I. HI-1. *Sol Phys*, 264:433–460, July 2010. doi: 10.1007/s11207-010-9582-8.

L. Biermann. Kometenschweife und solare Korpuskularstrahlung. *Zeitschrift fuer Astrophysik*, 29:274, 1951.

D.E. Billings. A guide to the solar corona: D.e. billings: Academic press, new york, 1966. 10+323. 112s. *Planetary and Space Science*, 15(12):1958 –, 1967. ISSN 0032-0633. doi: http://dx.doi.org/10.1016/0032-0633(67)90030-X.

D. S. Brown, D. Bewsher, and C. J. Eyles. Calibrating the Pointing and Optical Parameters of the STEREO Heliospheric Imagers. *Sol Phys*, 254:185–225, January 2009. doi: 10.1007/s11207-008-9277-6.

G. E. Brueckner, R. A. Howard, M. J. Koomen, C. M. Korendyke, D. J. Michels, J. D. Moses, D. G. Socker, K. P. Dere, P. L. Lamy, A. Llebaria, M. V. Bout, R. Schwenn, G. M. Simnett, D. K. Bedford, and C. J. Eyles. The Large Angle Spectroscopic Coronagraph (LASCO). *Sol Phys*, 162:357–402, December 1995. doi: 10.1007/BF00733434.

J. T. Burkepile and O. C. St. Cyr. A revised and expanded catalogue of mass ejections observed by the Solar Maximum Mission coronagraph. *NASA STI/Recon Technical Report N*, 93, January 1993.

A. G. Burns, W. Wang, M. Wiltberger, S. C. Solomon, H. Spence, T. L. Killeen, R. E. Lopez, and J. E. Landivar. An event study to provide validation of ting and cmit geomagnetic middle-latitude electron densities at the f2 peak. *Journal of Geophysical Research: Space Physics*, 113(A5):n/a–n/a, 2008. ISSN 2156-2202. doi: 10.1029/2007JA012931. URL `http://dx.doi.org/10.1029/2007JA012931`. A05310.

M. D. Butala, R. J. Hewett, R. A. Frazin, and F. Kamalabadi. Dynamic Three-Dimensional Tomography of the Solar Corona. *Sol Phys*, 262:495–509, April 2010. doi: 10.1007/s11207-010-9536-1.

P. J. Cargill. On the Aerodynamic Drag Force Acting on Interplanetary Coronal Mass Ejections. *Sol Phys*, 221:135–149, May 2004. doi: 10.1023/B:SOLA.0000033366.10725.a2.

R. C. Carrington. Description of a singular appearance seen in the sun on september 1, 1859. *R. Astron. Soc.*, XX:13, 1859.

S. Chapman and R.S. Lindzen. *Atmospheric Tides: Thermal and Gravitational.* Astrophysics and space science library. Reidel, 1970. ISBN 9789027701138. URL `http://books.google.co.uk/books?id=fS\_TJ63wdAYC`.

G. Chisham, M. Lester, S. Milan, M. Freeman, W. Bristow, A. Grocott, K. McWilliams, J. Ruohoniemi, T. Yeoman, P. Dyson, R. Greenwald, T. Kikuchi, M. Pinnock, J. Rash, N. Sato, G. Sofko, J.-P. Villain, and A. Walker. A decade of the super dual auroral radar network (superdarn): scientific achievements, new techniques and future directions. *Surveys in Geophysics*, 28:33–109, 2007. ISSN 0169-3298. 10.1007/s10712-007-9017-8.

M. V. Codrescu, T. J. Fuller-Rowell, and J. C. Foster. On the importance of E-field variability for Joule heating in the high-latitude thermosphere. *Geophysical Research Letters*, 22:2393–2396, 1995. doi: 10.1029/95GL01909.

M. V. Codrescu, T. J. Fuller-Rowell, J. C. Foster, J. M. Holt, and S. J. Cariglia. Electric field variability associated with the Millstone Hill electric field model. *Journal of Geophysical Research*, 105:5265–5274, March 2000. doi: 10.1029/1999JA900463.

F. D. Colegrove, F. S. Johnson, and W. B. Hanson. Atmospheric composition in the lower thermosphere. *Journal of Geophysical Research*, 71(9):2227–2236, 1966. ISSN 2156-2202. doi: 10.1029/JZ071i009p02227.

E. D. P. Cousins and S. G. Shepherd. A dynamical model of high-latitude convection derived from superdarn plasma drift measurements. *Journal of Geophysical Research: Space Physics*, 115(A12):n/a–n/a, 2010. ISSN 2156-2202. doi: 10.1029/2010JA016017. URL `http://dx.doi.org/10.1029/2010JA016017`.

A. Dalgarno and F.J. Smith. The thermal conductivity and viscosity of atomic oxygen. *Planetary and Space Science*, 9(12):1 – 2, 1962. ISSN 0032-0633. doi: 10.1016/0032-0633(62)90064-8.

J. A. Davies, R. A. Harrison, A. P. Rouillard, N. R. Sheeley, C. H. Perry, D. Bewsher, C. J. Davis, C. J. Eyles, S. R. Crothers, and D. S. Brown. A synoptic view of solar transient evolution in the inner heliosphere using the Heliospheric Imagers on STEREO. *Geophysical Research Letters*, 36:L02102, January 2009. doi: 10.1029/2008GL036182.

J. A. Davies, R. A. Harrison, C. H. Perry, C. Mstl, N. Lugaz, T. Rollett, C. J. Davis, S. R. Crothers, M. Temmer, C. J. Eyles, and N. P. Savani. A self-similar expansion model for use in solar wind transient propagation studies. *The Astrophysical Journal*, 750(1):23, 2012. URL `http://stacks.iop.org/0004-637X/750/i=1/a=23`.

J. A. Davies, C. H. Perry, R. M. G. M. Trines, R. A. Harrison, N. Lugaz, C. Möstl, Y. D. Liu, and K. Steed. Establishing a Stereoscopic Technique for Determining the Kinematic Properties of Solar Wind Transients based on a Generalized Self-similarly Expanding Circular Geometry. *ApJ*, 777:167, November 2013. doi: 10.1088/0004-637X/777/2/167.

J. M. Davila. Solar Tomography. *ApJ*, 423:871, March 1994. doi: 10.1086/173864.

J. M. Davila, D. M. Rust, V. J. Pizzo, and P. C. Liewer. Solar Terrestrial Relations Observatory (STEREO). In D. M. Rust, editor, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 2804 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 34–38, November 1996.

C. J. Davis, J. A. Davies, M. Lockwood, A. P. Rouillard, C. J. Eyles, and R. A. Harrison. Stereoscopic imaging of an Earth-impacting solar coronal mass ejection: A major milestone for the STEREO mission. *Geophysical Research Letters*, 36:L08102, April 2009. doi: 10.1029/2009GL038021.

C. J. Davis, J. A. Davies, O. C. St Cyr, M. Campbell-Brown, A. Skelt, M. Kaiser, N. Meyer-Vernet, S. Crothers, C. Lintott, A. Smith, S. Bamford, and E. M. L. Baeten. The distribution of interplanetary dust between 0.96 and 1.04 au as inferred from impacts on the STEREO spacecraft observed by the heliospheric imagers. *MNRAS*, 420:1355–1366, February 2012. doi: 10.1111/j.1365-2966.2011.20125.x.

J. de Patoul, B. Inhester, L. Feng, and T. Wiegelmann. 2D and 3D Polar Plume Analysis from the Three Vantage Positions of STEREO/EUVI A, B, and SOHO/EIT. *Sol Phys*, 283:207–225, March 2013. doi: 10.1007/s11207-011-9902-7.

Yue Deng and Aaron J. Ridley. Possible reasons for underestimating joule heating in global models: E field variability, spatial resolution, and vertical velocity. *Journal of Geophysical Research: Space Physics*, 112 (A9):n/a–n/a, 2007. ISSN 2156-2202. doi: 10.1029/2006JA012006. URL `http://dx.doi.org/10.1029/2006JA012006`.

A. L. Dobbin. Modelling studies of possible coupling mechanisms between the upper and middle atmosphere. *University College London*, PhD. Thesis, 2005.

A. Driesman, S. Hynes, and G. Cancro. The STEREO Observatory. *Space Science Reviews*, 136:17–44, April 2008. doi: 10.1007/s11214-007-9286-z.

B. a. Emery, C. Lathuillere, P. G. Richards, R. G. Roble, M. J. Buonsanto, D. J. Knipp, P. Wilkinson, D. P. Sipler, and R. Niciejewski. Time dependent thermospheric neutral response tothe 2-11 November 1993 storm period. *Journal of Atmospheric and Solar-Terrestrial Physics*, 61:329–350, February 1999. doi: 10.1016/S1364-6826(98)00137-0.

C. J. Eyles, G. M. Simnett, M. P. Cooke, B. V. Jackson, A. Buffington, P. P. Hick, N. R. Waltham, J. M. King, P. A. Anderson, and P. E. Holladay. The Solar Mass Ejection Imager (Smei). *Sol Phys*, 217: 319–347, November 2003. doi: 10.1023/B:SOLA.0000006903.75671.49.

C. J. Eyles, R. A. Harrison, C. J. Davis, N. R. Waltham, B. M. Shaughnessy, H. C. A. Mapson-Menard, D. Bewsher, S. R. Crothers, J. A. Davies, G. M. Simnett, R. A. Howard, J. D. Moses, J. S. Newmark, D. G. Socker, J.-P. Halain, J.-M. Defise, E. Mazy, and P. Rochus. The Heliospheric Imagers Onboard the STEREO Mission. *Sol Phys*, 254:387–445, February 2009. doi: 10.1007/s11207-008-9299-0.

Rohantha P. Fernando and Ian W.M. Smith. Vibrational relaxations of {NO} by atomic oxygen. *Chemical Physics Letters*, 66(2):218 – 222, 1979. ISSN 0009-2614. doi: 10.1016/0009-2614(79)85002-2.

C. C. Finlay, S. Maus, C. D. Beggan, T. N. Bondar, A. Chambodut, T. A. Chernova, A. Chulliat, V. P. Golovkov, B. Hamilton, M. Hamoudi, R. Holme, G. Hulot, W. Kuang, B. Langlais, V. Lesur, F. J. Lowes, H. Lühr, S. MacMillan, M. Mandea, S. McLean, C. Manoj, M. Menvielle, I. Michaelis, N. Olsen, J. Rauberg, M. Rother, T. J. Sabaka, A. Tangborn, L. Tøffner-Clausen, E. Thébault, A. W. P. Thomson, I. Wardinski, Z. Wei, and T. I. Zvereva. International Geomagnetic Reference Field: the eleventh generation. *Geophysical Journal International*, 183:1216–1230, December 2010. doi: 10.1111/j.1365-246X.2010.04804.x.

R. R. Fisher, R. H. Lee, R. M. MacQueen, and A. I. Poland. New Mauna Loa coronagraph systems. *Applied Optics*, 20:1094–1101, March 1981. doi: 10.1364/AO.20.001094.

V. I. Fomichev, J.-P. Blanchet, and D. S. Turner. Matrix parameterization of the 15 m co2 band cooling in the middle and upper atmosphere for variable co2 concentration. *Journal of Geophysical Research: Atmospheres*, 103(D10):11505–11528, 1998. ISSN 2156-2202. doi: 10.1029/98JD00799.

M. Förster and I. Cnossen. Upper atmosphere differences between northern and southern high latitudes: The role of magnetic field asymmetry. *Journal of Geophysical Research (Space Physics)*, 118:5951–5966, September 2013. doi: 10.1002/jgra.50554.

M. Förster, S. E. Haaland, S. Rentz, and H. Liu. IMF Dependence of High-Latitude Thermospheric Wind Pattern Derived from CHAMP Cross-Track Accelerometer Data and the Corresponding Magnetospheric Convection from Cluster EDI Measurements. In *37th COSPAR Scientific Assembly*, volume 37 of *COSPAR Meeting*, page 902, 2008.

M. Förster, B. E. Prokhorov, A. A. Namgaladze, and M. Holschneider. Numerical modeling of solar wind influences on the dynamics of the high-latitude upper atmosphere. *Advances in Radio Science*, 10: 299–312, 2012. doi: 10.5194/ars-10-299-2012. URL `http://www.adv-radio-sci.net/10/299/2012/`.

J. C. Foster, J. M. Holt, R. G. Musgrove, and D. S. Evans. Ionospheric convection associated with discrete levels of particle precipitation. *Geophysical Research Letters*, 13:656–659, July 1986. doi: 10.1029/GL013i007p00656.

A. C. Fraser-Smith. Centered and eccentric geomagnetic dipoles and their poles, 1600 - 1985. *Reviews of Geophysics*, 25:1–16, 1987. doi: 10.1029/RG025i001p00001.

R. A. Frazin. Tomography of the Solar Corona. I. A Robust, Regularized, Positive Estimation Method. *ApJ*, 530:1026–1035, February 2000. doi: 10.1086/308412.

R. A. Frazin. Coronal Mass Ejection Reconstruction from Three Viewpoints via Simulation Morphing. I. Theory and Examples. *ApJ*, 761:24, December 2012. doi: 10.1088/0004-637X/761/1/24.

R. A. Frazin and P. Janzen. Tomography of the Solar Corona. II. Robust, Regularized, Positive Estimation of the Three-dimensional Electron Density Distribution from LASCO-C2 Polarized White-Light Images. *ApJ*, 570:408–422, May 2002. doi: 10.1086/339572.

R. A. Frazin, M. Jacob, W. B. Manchester IV, H. Morgan, and M. B. Wakin. Toward reconstruction of coronal mass ejection density from only three points of view. *The Astrophysical Journal*, 695(1):636, 2009. URL `http://stacks.iop.org/0004-637X/695/i=1/a=636`.

R. A. Frazin, M. Jacob, W. B. Manchester, IV, H. Morgan, and M. B. Wakin. Toward Reconstruction of Coronal Mass Ejection Density from Only Three Points of View. *ApJ*, 695:636–641, April 2009. doi: 10.1088/0004-637X/695/1/636.

T. J. Fuller-Rowell. A three-dimensional time-dependant global model of the thermosphere. *University College London*, PhD. Thesis, 1981.

T. J. Fuller-Rowell. A two-dimensional, high-resolution, nested-grid model of the thermosphere: 1. neutral response to an electric field spike. *Journal of Geophysical Research: Space Physics*, 89(A5):2971–2990, 1984. ISSN 2156-2202. doi: 10.1029/JA089iA05p02971.

T. J. Fuller-Rowell and D. S. Evans. Height-integrated Pedersen and Hall conductivity patterns inferred from the TIROS-NOAA satellite data. *Journal of Geophysical Research*, 92:7606–7618, July 1987. doi: 10.1029/JA092iA07p07606.

T. J. Fuller-Rowell and D. Rees. A three-dimensional time-dependent global model of the thermosphere. *Journal of Atmospheric Sciences*, 37:2545–2567, nov 1980. doi: 10.1175/1520-0469(1980)037⟨2545:ATDTDG⟩2.0.CO;2.

T. J. Fuller-Rowell and D. Rees. Derivation of a conservation equation for mean molecular weight for a two-constituent gas within a three-dimensional time-dependent model of the thermosphere. *Planetary Space Science*, 31:1209–1222, October 1983. doi: 10.1016/0032-0633(83)90112-5.

Rolando R. Garcia and Susan Solomon. A numerical model of the zonally averaged dynamical and chemical structure of the middle atmosphere. *Journal of Geophysical Research: Oceans*, 88(C2):1379–1400, 1983. ISSN 2156-2202. doi: 10.1029/JC088iC02p01379.

P. R. Gazis, A. Balogh, S. Dalla, R. Decker, B. Heber, T. Horbury, A. Kilchenmann, J. Kota, H. Kucharek, H. Kunow, D. Lario, M. S. Potgieter, J. D. Richardson, P. Riley, L. Rodriguez, G. Siscoe, and R. von Steiger. ICMEs at High Latitudes and in the Outer Heliosphere. Report of Working Group H. *Space Science Reviews*, 123:417–451, March 2006. doi: 10.1007/s11214-006-9023-z.

R. G. Gillies, G. C. Hussey, G. J. Sofko, P. V. Ponomarenko, and K. A. McWilliams. Improvement of HF coherent radar line-of-sight velocities by estimating the refractive index in the scattering volume using radar frequency shifting. *Journal of Geophysical Research (Space Physics)*, 116:A01302, January 2011. doi: 10.1029/2010JA016043.

R. G. Gillies, G. C. Hussey, G. J. Sofko, and K. A. McWilliams. A statistical analysis of SuperDARN scattering volume electron densities and velocity corrections using a radar frequency shifting technique. *Journal of Geophysical Research (Space Physics)*, 117:A08320, August 2012. doi: 10.1029/2012JA017866.

W. D. Gonzalez, J. A. Joselyn, Y. Kamide, H. W. Kroehl, G. Rostoker, B. T. Tsurutani, and V. M. Vasyliunas. What is a geomagnetic storm? *Journal of Geophysical Research*, 99:5771–5792, April 1994. doi: 10.1029/93JA02867.

N Gopalswamy, S Yashiro, G Michalek, G Stenborg, A Vourlidas, S Freeland, and R Howard. The soho/lasco cme catalog. *Earth, Moon, and Planets*, 104(1):295–313, 2009.

J. T. Gosling. The solar flare myth. *Journal of Geophysical Research*, 98:18937–18950, November 1993. doi: 10.1029/93JA01896.

J. T. Gosling, E. Hildner, R. M. MacQueen, R. H. Munro, A. I. Poland, and C. L. Ross. Mass ejections from the sun - A view from SKYLAB. *Journal of Geophysical Research*, 79:4581–4587, November 1974. doi: 10.1029/JA079i031p04581.

R. A. Greenwald, K. B. Baker, J. R. Dudeney, M. Pinnock, T. B. Jones, E. C. Thomas, J.-P. Villain, J.-C. Cerisier, C. Senior, C. Hanuise, R. D. Hunsucker, G. Sofko, J. Koehler, E. Nielsen, R. Pellinen, A. D. M. Walker, N. Sato, and H. Yamagishi. Darn/Superdarn: A Global View of the Dynamics of High-Lattitude Convection. *Space Science Reviews*, 71:761–796, February 1995. doi: 10.1007/BF00751350.

A. Grocott and S. E. Milan. The influence of IMF clock angle timescales on the morphology of ionospheric convection. *Journal of Geophysical Research (Space Physics)*, 119:5861–5876, July 2014. doi: 10.1002/2014JA020136.

J.-P. Halain, C. J. Eyles, A. Mazzoli, D. Bewsher, J. A. Davies, E. Mazy, P. Rochus, J. M. Defise, C. J. Davis, R. A. Harrison, S. R. Crothers, D. S. Brown, C. Korendyke, J. D. Moses, D. G. Socker, R. A. Howard, and J. S. Newmark. Straylight-Rejection Performance of the STEREO HI Instruments. *Sol Phys*, 271:197–218, July 2011. doi: 10.1007/s11207-011-9800-z.

M. J. Harris. A new coupled middle atmosphere and thermosphere circulation model: Studies of dynamic, energetic and photochemical coupling in the middle and upper atmosphere. *University College London*, PhD. Thesis, 2001.

J. H. Hecht, A. B. Christensen, J. B. Pranke, W. T. Chater, C. K. Howey, R. L. Lott, and M. G. Sivjee. Auroral and airglow fabryperot spectrometer. *Review of Scientific Instruments*, 57(2), 1986.

A. E. Hedin. Extension of the MSIS thermosphere model into the middle and lower atmosphere. *Journal of Geophysical Research*, 96:1159–1172, February 1991. doi: 10.1029/90JA02125.

A. E. Hedin, C. A. Reber, G. P. Newton, N. W. Spencer, H. C. Brinton, H. G. Mayr, and W. E. Potter. A global thermospheric model based on mass spectrometer and incoherent scatter data MSIS. II - Composition. *Journal of Geophysical Research*, 82:2148–2156, June 1977a. doi: 10.1029/JA082i016p02148.

A. E. Hedin, C. A. Reber, G. P. Newton, N. W. Spencer, J. E. Salah, J. V. Evans, D. C. Kayser, D. Alcayde, P. Bauer, and L. Cogger. A global thermospheric model based on mass spectrometer and incoherent scatter data MSIS. I - N2 density and temperature. *Journal of Geophysical Research*, 82:2139–2147, June 1977b. doi: 10.1029/JA082i016p02139.

J. P. Heppner. Polar-cap electric field distributions related to the interplanetary magnetic field direction. *Journal of Geophysical Research*, 77(25):4877–4887, 1972. ISSN 2156-2202. doi: 10.1029/JA077i025p04877. URL `http://dx.doi.org/10.1029/JA077i025p04877`.

J. P. Heppner. Electric field variations during substorms: OGO-6 measurements. *Planetary Space Science*, 20:1475–1498, September 1972. doi: 10.1016/0032-0633(72)90052-9.

J. P. Heppner. Empirical models of high-latitude electric fields. *Journal of Geophysical Research*, 82(7): 1115–1125, 1977. ISSN 2156-2202. doi: 10.1029/JA082i007p01115. URL `http://dx.doi.org/10.1029/JA082i007p01115`.

J. P. Heppner and N. C. Maynard. Empirical high-latitude electric field models. *Journal of Geophysical Research*, 92:4467–4489, May 1987. doi: 10.1029/JA092iA05p04467.

Magnus R. Hestenes and Eduard Stiefel. Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, 49(6):409–436, December 1952.

J. R. Holton. The role of gravity wave induced drag and diffusion in the momentum budget of the mesosphere. *Journal of the Atmospheric Sciences*, 39, 1981. ISSN 0022-4928. doi: 10.1029/JC086iC10p09707. URL `http://dx.doi.org/10.1175/1520-0469(1982)039<0791:TROGWI>2.0.CO;2`.

J. T. Houghton. The stratosphere and mesosphere. *Quarterly Journal of the Royal Meteorological Society*, 104(439):1–29, 1978. ISSN 1477-870X. doi: 10.1002/qj.49710443902. URL `http://dx.doi.org/10.1002/qj.49710443902`.

L. L. House, W. J. Wagner, E. Hildner, C. Sawyer, and H. U. Schmidt. Studies of the corona with the Solar Maximum Mission coronagraph/polarimeter. *ApJL*, 244:L117–L121, March 1981. doi: 10.1086/183494.

R. A. Howard, D. J. Michels, N. R. Sheeley, Jr., and M. J. Koomen. The observation of a coronal transient directed at earth. *ApJL*, 263:L101–L104, December 1982. doi: 10.1086/183932.

R. A. Howard, N. R. Sheeley, Jr., D. J. Michels, and M. J. Koomen. Coronal mass ejections - 1979-1981. *Journal of Geophysical Research*, 90:8173–8191, September 1985. doi: 10.1029/JA090iA09p08173.

R. A. Howard, J. D. Moses, A. Vourlidas, J. S. Newmark, D. G. Socker, S. P. Plunkett, C. M. Korendyke, J. W. Cook, A. Hurley, J. M. Davila, W. T. Thompson, O. C. St Cyr, E. Mentzell, K. Mehalick, J. R. Lemen, J. P. Wuelser, D. W. Duncan, T. D. Tarbell, C. J. Wolfson, A. Moore, R. A. Harrison, N. R. Waltham, J. Lang, C. J. Davis, C. J. Eyles, H. Mapson-Menard, G. M. Simnett, J. P. Halain, J. M. Defise, E. Mazy, P. Rochus, R. Mercier, M. F. Ravet, F. Delmotte, F. Auchere, J. P. Delaboudiniere, V. Bothmer, W. Deutsch, D. Wang, N. Rich, S. Cooper, V. Stephens, G. Maahs, R. Baugh, D. McMullin, and T. Carter. Sun Earth Connection Coronal and Heliospheric Investigation (SECCHI). *Space Science Reviews*, 136:67–115, April 2008. doi: 10.1007/s11214-008-9341-4.

S. M. Imber, S. E. Milan, and M. Lester. The Heppner-Maynard Boundary measured by SuperDARN as a proxy for the latitude of the auroral oval. *Journal of Geophysical Research (Space Physics)*, 118:685–697, February 2013. doi: 10.1029/2012JA018222.

B. V. Jackson and H. R. Froehling. Three-dimensional reconstruction of a coronal mass ejection. *A&A*, 299:885, July 1995.

B. V. Jackson, P. P. Hick, and A. Buffington. Time-dependent tomography of heliospheric structures using IPS and Thomson scattering observations. In A. Wilson, editor, *Solar Variability as an Input to the Earth's Environment*, volume 535 of *ESA Special Publication*, pages 823–833, September 2003.

J. D. Jackson. *Classical electrodynamics*. 1975.

P. Janhunen. GUMICS-3 A Global Ionosphere-Magnetosphere Coupling Simulation with High Ionospheric Resolution. In T.-D. Guyenne and A. Hilgers, editors, *Environment Modeling for Space-Based Applications*, volume 392 of *ESA Special Publication*, page 233, December 1996.

P. Janhunen. The electric solar wind sail status report. In *European Planetary Science Congress 2010*, page 297, September 2010.

J. R. Jokipii and B. Thomas. Effects of drift on the transport of cosmic rays. IV - Modulation by a wavy interplanetary current sheet. *ApJ*, 243:1115–1122, February 1981. doi: 10.1086/158675.

R.A. Jones and M.H. Rees. Time dependent studies of the aurora. ion density and composition. *Planetary and Space Science*, 21(4):537 – 557, 1973. ISSN 0032-0633. doi: 10.1016/0032-0633(73)90069-X.

M. L. Kaiser, T. A. Kucera, J. M. Davila, O. C. St. Cyr, M. Guhathakurta, and E. Christian. The STEREO Mission: An Introduction. *Space Science Reviews*, 136:5–16, April 2008. doi: 10.1007/s11214-007-9277-0.

Y. Kamide, A. D. Richmond, and S. Matsushita. Estimation of ionospheric electric fields, ionospheric currents, and field-aligned currents from ground magnetic records. *Journal of Geophysical Research*, 86: 801–813, February 1981. doi: 10.1029/JA086iA02p00801.

J. R. Kan and L. C. Lee. Energy coupling function and solar wind-magnetosphere dynamo. *Geophysical Research Letters*, 6:577–580, July 1979. doi: 10.1029/GL006i007p00577.

S. R. Kane, K. Hurley, J. M. McTiernan, M. Sommer, M. Boer, and M. Niel. Energy Release and Dissipation during Giant Solar Flares. *ApJL*, 446:L47, June 1995. doi: 10.1086/187927.

H. Khan and S. W. H. Cowley. Observations of the response time of high-latitude ionospheric convection to variations in the interplanetary magnetic field using EISCAT and IMP-8 data. *Annales Geophysicae*, 17:1306–1335, October 1999. doi: 10.1007/s00585-999-1306-8.

D. J. Knipp, A. D. Richmond, B. Emery, N. U. Crooker, O. de La Beaujardiere, and D. Evans. Ionospheric convection response to changing IMF direction. *Geophysical Research Letters*, 18:721–724, April 1991. doi: 10.1029/90GL02592.

D. J. Knipp, W. K. Tobiska, and B. Emery. Direct and Indirect Thermospheric Heating Sources. *AGU Spring Meeting Abstracts*, page A1, May 2005.

G. Kockarts. Nitric oxide cooling in the terrestrial thermosphere. *Geophysical Research Letters*, 7:137–140, February 1980. doi: 10.1029/GL007i002p00137.

M. Kramar, S. Jones, J. Davila, B. Inhester, and M. Mierla. On the Tomographic Reconstruction of the 3D Electron Density for the Solar Corona from STEREO COR1 Data. *Sol Phys*, 259:109–121, October 2009. doi: 10.1007/s11207-009-9401-2.

M. Kramar, V. Airapetian, Z. Mikić, and J. Davila. 3D Coronal Density Reconstruction and Retrieving the Magnetic Field Structure during Solar Minimum. *Sol Phys*, April 2014. doi: 10.1007/s11207-014-0525-7.

M. Kramar, H. Lin, and S. Tomczyk. Direct Observation of Solar Coronal Magnetic Fields by Vector Tomography of the Coronal Emission Line Polarizations. *ApJL*, 819:L36, March 2016. doi: 10.3847/2041-8205/819/2/L36.

S. V. Leontyev and W. B. Lyatsky. Electric field and currents connected with Y-component of interplanetary magnetic field. *Planetary Space Science*, 22:811–819, May 1974. doi: 10.1016/0032-0633(74)90151-2.

R. P. Lin and H. S. Hudson. Non-thermal processes in large solar flares. *Sol Phys*, 50:153–178, October 1976. doi: 10.1007/BF00206199.

R. S. Lindzen. Thermally driven diurnal tide in the atmosphere. *Quarterly Journal of the Royal Meteorological Society*, 93(395):18–42, 1967. ISSN 1477-870X. doi: 10.1002/qj.49709339503. URL `http://dx.doi.org/10.1002/qj.49709339503`.

R. S. Lindzen. Turbulence and stress owing to gravity wave and tidal breakdown. *Journal of Geophysical Research: Oceans*, 86(C10):9707–9714, 1981. ISSN 2156-2202. doi: 10.1029/JC086iC10p09707. URL `http://dx.doi.org/10.1029/JC086iC10p09707`.

H. Liu, H. Lühr, V. Henize, and W. KöHler. Global distribution of the thermospheric total mass density derived from CHAMP. *Journal of Geophysical Research (Space Physics)*, 110:A04301, April 2005. doi: 10.1029/2004JA010741.

B. C. Low. Coronal mass ejections, magnetic flux ropes, and solar magnetism. *Journal of Geophysical Research*, 106:25141–25164, November 2001. doi: 10.1029/2000JA004015.

G. Lu, A. D. Richmond, B. A. Emery, P. H. Reiff, O. de La Beaujardiere, F. J. Rich, W. F. Denig, H. W. Kroehl, L. R. Lyons, and J. M. Ruohoiemi. Interhemispheric asymmetry of the high-latitude ionospheric convection pattern. *Journal of Geophysical Research*, 99:6491–6510, April 1994. doi: 10.1029/93JA03441.

A. T. Y. Lui. Current disruption in the Earth's magnetosphere: Observations and models. *Journal of Geophysical Research*, 101:13067–13088, June 1996. doi: 10.1029/96JA00079.

A. T. Y. Lui. Current controversies in magnetospheric physics. *Reviews of Geophysics*, 39(4):535–563, 2001. ISSN 1944-9208. doi: 10.1029/2000RG000090. URL `http://dx.doi.org/10.1029/2000RG000090`.

Peter Lynch. The origins of computer weather prediction and climate modeling. *Journal of Computational Physics*, 227(7):3431 – 3444, 2008. ISSN 0021-9991. doi: http://dx.doi.org/10.1016/j.jcp.2007.02.034. URL `http://www.sciencedirect.com/science/article/pii/S0021999107000952`. ¡ce:title¿Predicting weather, climate and extreme events¡/ce:title¿.

J. G. Lyon, J. A. Fedder, and C. M. Mobarry. The Lyon-Fedder-Mobarry (LFM) global MHD magnetospheric simulation code. *Journal of Atmospheric and Solar-Terrestrial Physics*, 66:1333–1350, October 2004. doi: 10.1016/j.jastp.2004.03.020.

R. M. MacQueen, A. Csoeke-Poeckh, E. Hildner, L. House, R. Reynolds, A. Stanger, H. Tepoel, and W. Wagner. The High Altitude Observatory Coronagraph/Polarimeter on the Solar Maximum Mission. *Sol Phys*, 65:91–107, February 1980. doi: 10.1007/BF00151386.

T. Matsuo and A. D. Richmond. Effects of high-latitude ionospheric electric field variability on global thermospheric Joule heating and mechanical energy transfer rate. *Journal of Geophysical Research (Space Physics)*, 113:A07309, July 2008. doi: 10.1029/2007JA012993.

S. L. McGregor, W. J. Hughes, C. N. Arge, D. Odstrcil, and N. A. Schwadron. The radial evolution of solar wind speeds. *Journal of Geophysical Research (Space Physics)*, 116:A03106, March 2011. doi: 10.1029/2010JA016006.

R. L. McPherron. Growth phase of magnetospheric substorms. *Journal of Geophysical Research*, 75:5592, 1970. doi: 10.1029/JA075i028p05592.

R. L. McPherron. Magnetospheric substorms. *Reviews of Geophysics and Space Physics*, 17:657–681, June 1979. doi: 10.1029/RG017i004p00657.

Zoran Miki, Jon A. Linker, Dalton D. Schnack, Roberto Lionello, and Alfonso Tarditi. Magnetohydrodynamic modeling of the global solar corona. *Physics of Plasmas*, 6(5), 1999.

S. E. Milan, M. Lester, S. W. H. Cowley, K. Oksavik, M. Brittnacher, R. A. Greenwald, G. Sofko, and J.-P. Villain. Variations in the polar cap area during two substorm cycles. *Annales Geophysicae*, 21: 1121–1140, May 2003. doi: 10.5194/angeo-21-1121-2003.

G. H. Millward, R. J. Moffett, S. Quegan, and T. J. Fuller-Rowell. A coupled thermosphere-ionosphere-plasmasphere model. *Solar Terrestrial Energy Program (STEP) Handbook*, 1996a.

M. Minnaert. On the continuous spectrum of the corona and its polarisation. With 3 figures. (Received July 30, 1930). *Z. Astrophys.*, 1:209, 1930.

T. G. Moran and J. M. Davila. Three-Dimensional Polarimetric Imaging of Coronal Mass Ejections. *Science*, 305:66–71, July 2004. doi: 10.1126/science.1098937.

C. Möstl, K. Amla, J. R. Hall, P. C. Liewer, E. M. De Jong, R. C. Colaninno, A. M. Veronig, T. Rollett, M. Temmer, V. Peinhart, J. A. Davies, N. Lugaz, Y. D. Liu, C. J. Farrugia, J. G. Luhmann, B. Vršnak, R. A. Harrison, and A. B. Galvin. Connecting speeds, directions and arrival times of 22 coronal mass ejections from the Sun to 1 AU. *ArXiv e-prints*, April 2014.

T. C. Mouschovias and A. I. Poland. Expansion and broadening of coronal loop transients - A theoretical explanation. *ApJ*, 220:675–682, March 1978. doi: 10.1086/155951.

D. O. Muhleman, P. B. Esposito, and J. D. Anderson. The electron density profile of the outer corona and the interplanetary medium from Mariner-6 and Mariner-7 time-delay measurements. *ApJ*, 211:943–957, February 1977. doi: 10.1086/155006.

R. J. Murgatroyd and F. Singleton. Possible meridional circulations in the stratosphere and mesosphere. *Quarterly Journal of the Royal Meteorological Society*, 87(372):125–135, 1961. ISSN 1477-870X. doi: 10.1002/qj.49708737202. URL `http://dx.doi.org/10.1002/qj.49708737202`.

D. Odstrcil and V. J. Pizzo. Distortion of the interplanetary magnetic field by three-dimensional propagation of coronal mass ejections in a structured solar wind. *Journal of Geophysical Research*, 104:28225–28240, 1999. doi: 10.1029/1999JA900319.

D. Odstrcil, V. J. Pizzo, J. A. Linker, P. Riley, R. Lionello, and Z. Mikic. Initial coupling of coronal and heliospheric numerical magnetohydrodynamic codes. *Journal of Atmospheric and Solar-Terrestrial Physics*, 66:1311–1320, October 2004. doi: 10.1016/j.jastp.2004.04.007.

V. O. Papitashvili and F. J. Rich. High-latitude ionospheric convection models derived from Defense Meteorological Satellite Program ion drift observations and parameterized by the interplanetary magnetic field strength and direction. *Journal of Geophysical Research (Space Physics)*, 107:1198, August 2002. doi: 10.1029/2001JA000264.

E. N. Parker. Dynamics of the Interplanetary Gas and Magnetic Fields. *ApJ*, 128:664, November 1958. doi: 10.1086/146579.

E. N. Parker. The Hydrodynamic Treatment of the Expanding Solar Corona. *ApJ*, 132:175, July 1960. doi: 10.1086/146910.

P. Perreault and S.-I. Akasofu. A study of geomagnetic storms. *Geophysical Journal International*, 54: 547–573, September 1978. doi: 10.1111/j.1365-246X.1978.tb05494.x.

J. M. Picone, A. E. Hedin, D. P. Drob, and A. C. Aikin. NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues. *Journal of Geophysical Research (Space Physics)*, 107:1468, December 2002. doi: 10.1029/2002JA009430.

V. J. Pizzo. A three-dimensional model of corotating streams in the solar wind. II - Hydrodynamic streams. *Journal of Geophysical Research*, 85:727–743, February 1980. doi: 10.1029/JA085iA02p00727.

G. W. Pneuman and R. A. Kopp. Gas-Magnetic Field Interactions in the Solar Corona. *Sol Phys*, 18: 258–270, June 1971. doi: 10.1007/BF00145940.

S Quegan, G.J Bailey, R.J Moffett, R.A Heelis, T.J Fuller-Rowell, D Rees, and R.W Spiro. A theoretical study of the distribution of ionization in the high-latitude ionosphere and the plasmasphere: first results on the mid-l. *Journal of Atmospheric and Terrestrial Physics*, 44(7):619 – 640, 1982. ISSN 0021-9169. doi: 10.1016/0021-9169(82)90073-3.

C. A. Reber. Thermospheric wind effects on the global distribution of helium in the earth's upper atmosphere. 1974.

H. Rees. *Physics and Chemistry of the Upper Atmosphere.* Cambridge Atmospheric and Space Science Series. Cambridge University Press, 1989. ISBN 9780521323055.

John D. Richardson, Ying Liu, and John W. Belcher. Propagation and evolution of icmes in the solar wind. In Jean-Andr Sauvaud and Zden0̌11bk N0̌11bme0̌10dek, editors, *Multiscale Processes in the Earth2̌019s Magnetosphere: From Interball to Cluster*, volume 178 of *NATO Science Series*, pages 1–14. Springer Netherlands, 2005. ISBN 978-1-4020-2768-0. 10.1007/1-4020-2768-0.

A. J. Ridley. A new formulation for the ionospheric cross polar cap potential including saturation effects. *Annales Geophysicae*, 23(11):3533–3547, 2005. doi: 10.5194/angeo-23-3533-2005.

A. J. Ridley, G. Lu, C. R. Clauer, and V. O. Papitashvili. A statistical study of the ionospheric convection response to changing interplanetary magnetic field conditions using the assimilative mapping of ionospheric electrodynamics technique. *Journal of Geophysical Research*, 103:4023–4040, March 1998. doi: 10.1029/97JA03328.

A. J. Ridley, Y. Deng, and G. Tóth. The global ionosphere thermosphere model. *Journal of Atmospheric and Solar-Terrestrial Physics*, 68:839–864, May 2006. doi: 10.1016/j.jastp.2006.01.008.

Y. Rippeth. Theoretical studies of ionosphere-plasmasphere coupling: the effects of an eccentric dipole geomagnetic field. *University of Sheffield*, PhD. Thesis, 1992.

H Rishbeth. Thermospheric winds and the f-region: A review. *Journal of Atmospheric and Terrestrial Physics*, 34(1):1 – 47, 1972. ISSN 0021-9169. doi: 10.1016/0021-9169(72)90003-7.

H. Rishbeth and O. K. Garriott. *Introduction to ionospheric physics.* 1969.

A. S. Rodger, G. D. Wells, R. J. Moffett, and G. J. Bailey. The variability of Joule heating, and its effects on the ionosphere and thermosphere. *Annales Geophysicae*, 19(7):773–781, 2001. URL `https://hal.archives-ouvertes.fr/hal-00316874`.

G. Rostoker, S.-I. Akasofu, J. Foster, R. A. Greenwald, A. T. Y. Lui, Y. Kamide, K. Kawasaki, R. L. McPherron, and C. T. Russell. Magnetospheric substorms - Definition and signatures. *Journal of Geophysical Research*, 85:1663–1668, April 1980. doi: 10.1029/JA085iA04p01663.

J. M. Ruohoniemi and K. B. Baker. Large-scale imaging of high-latitude convection with Super Dual Auroral Radar Network HF radar observations. *Journal of Geophysical Research*, 103:20797–20811, September 1998. doi: 10.1029/98JA01288.

J. M. Ruohoniemi and R. A. Greenwald. Statistical patterns of high-latitude convection obtained from Goose Bay HF radar observations. *Journal of Geophysical Research*, 101:21743–21764, October 1996. doi: 10.1029/96JA01584.

C. T. Russell, G. Lu, and J. G. Luhmann. Lessons from the ring current injection during the September 24, 25, 1998 storm. *Geophysical Research Letters*, 27:1371–1374, 2000. doi: 10.1029/1999GL003718.

D. M. Rust. *The Solar Stereo Mission*, pages 213–224. American Geophysical Union, 1999. ISBN 9781118664469. doi: 10.1029/GM109p0213. URL `http://dx.doi.org/10.1029/GM109p0213`.

Joseph E. Salah. Interim standard for the ion-neutral atomic oxygen collision frequency. *Geophysical Research Letters*, 20(15):1543–1546, 1993. ISSN 1944-8007. doi: 10.1029/93GL01699.

W.K.H Schmidt and V Bothmer. Stereoscopic viewing of solar coronal and interplanetary activity. *Advances in Space Research*, 17(45):369 – 376, 1996. ISSN 0273-1177. doi: http://dx.doi.org/10.1016/0273-1177(95)00602-B. URL `http://www.sciencedirect.com/science/article/pii/027311779500602B`. ¡ce:title¿Solar Flare, Coronal and Heliospheric Dynamics¡/ce:title¿.

R.W. Schunk and J.C.G. Walker. Theoretical ion densities in the lower ionosphere. *Planetary and Space Science*, 21(11):1875 – 1896, 1973. ISSN 0032-0633. doi: 10.1016/0032-0633(73)90118-9.

N. R. Sheeley, J. H. Walters, Y.-M. Wang, and R. A. Howard. Continuous tracking of coronal outflows: Two kinds of coronal mass ejections. *Journal of Geophysical Research*, 104:24739–24768, November 1999. doi: 10.1029/1999JA900308.

N. R. Sheeley, Jr., D. J. Michels, R. A. Howard, and M. J. Koomen. Initial observations with the SOLWIND coronagraph. *ApJ*, 237:L99–L101, May 1980. doi: 10.1086/183243.

K. Shiokawa, G. Lu, Y. Otsuka, T. Ogawa, M. Yamamoto, N. Nishitani, and N. Sato. Ground observation and amie-tiegcm modeling of a storm-time traveling ionospheric disturbance. *Journal of Geophysical Research: Space Physics*, 112(A5):n/a–n/a, 2007. ISSN 2156-2202. doi: 10.1029/2006JA011772. URL `http://dx.doi.org/10.1029/2006JA011772`. A05308.

G. L. Siscoe, N. U. Crooker, and K. D. Siebert. Transpolar potential saturation: Roles of region 1 current system and solar wind ram pressure. *Journal of Geophysical Research (Space Physics)*, 107:1321, October 2002. doi: 10.1029/2001JA009176.

S Solomon, R Garcia, J Olivero, R Bevilacqua, P Schwartz, R Clancy, and D Muhleman. Photochemistry and transport of carbon-monoxide in the middle atmosphere. 42(10), 1985.

J. Spaleta, W. A. Bristow, and J. Klein. Temporal and spatial resolved SuperDARN line of sight velocity measurements corrected for plasma index of refraction using Bayesian inference. *Journal of Geophysical Research (Space Physics)*, 120:3207–3225, April 2015. doi: 10.1002/2014JA020960.

Friedhelm Steinhilber, Jose A. Abreu, Jrg Beer, Irene Brunner, Marcus Christl, Hubertus Fischer, Ulla Heikkil, Peter W. Kubik, Mathias Mann, Ken G. McCracken, Heinrich Miller, Hiroko Miyahara, Hans Oerter, and Frank Wilhelms. 9,400 years of cosmic radiation and solar activity from ice cores and tree rings. *Proceedings of the National Academy of Sciences*, 109(16):5967–5971, 2012. doi: 10.1073/pnas.1118965109. URL `http://www.pnas.org/content/109/16/5967.abstract`.

W. Swider and M.E. Gardner. *On the Accuracy of Certain Approximations for the Chapman Function*. Environmental Research Papers. United States Air Force, Office of Aerospace Research, Air Force Cambridge Research Laboratories, Upper Atmosphere Physics Laboratory, 1967.

S. J. Tappin. The Deceleration of an Interplanetary Transient from the Sun to 5 Au. *Sol Phys*, 233:233–248, February 2006. doi: 10.1007/s11207-006-2065-2.

S. J. Tappin and T. A. Howard. Interplanetary Coronal Mass Ejections Observed in the Heliosphere: 1. Review of Theory. *Space Science Reviews*, 147:31–54, October 2009a. doi: 10.1007/s11214-009-9542-5.

S. J. Tappin and T. A. Howard. Interplanetary Coronal Mass Ejections Observed in the Heliosphere: 2. Model and Data Comparison. *Space Science Reviews*, 147:55–87, October 2009b. doi: 10.1007/s11214-009-9550-5.

S. J. Tappin and T. A. Howard. Interplanetary Coronal Mass Ejections Observed in the Heliosphere: 3. Physical Implications. *Space Science Reviews*, 147:89–110, October 2009c. doi: 10.1007/s11214-009-9577-7.

J. P. Thayer. Height-resolved joule heating rates in the high-latitude e region and the influence of neutral winds. *Journal of Geophysical Research: Space Physics*, 103(A1):471–487, 1998. ISSN 2156-2202. doi: 10.1029/97JA02536. URL `http://dx.doi.org/10.1029/97JA02536`.

J. P. Thayer and T. L. Killeen. A kinematic analysis of the high-latitude thermospheric neutral circulation pattern. *Journal of Geophysical Research*, 98:11549, July 1993. doi: 10.1029/93JA00629.

M. R. Torr, D. G. Torr, R. A. Ong, and H. E. Hinteregger. Ionization frequencies for major thermospheric constituents as a function of solar cycle 21. *Geophysical Research Letters*, 6:771–774, October 1979. doi: 10.1029/GL006i010p00771.

R. Tousey, R. A. Howard, and M. J. Koomen. The Frequency and Nature of Coronal Transient Events Observed by OSO-7*. In *Bulletin of the American Astronomical Society*, volume 6 of *Bulletin of the American Astronomical Society*, page 295, March 1974.

B. T. Tsurutani, W. D. Gonzalez, G. S. Lakhina, and S. Alex. The extreme magnetic storm of 12 september 1859. *Journal of Geophysical Research: Space Physics*, 108(A7):n/a–n/a, 2003. ISSN 2156-2202. doi: 10.1029/2002JA009504. URL `http://dx.doi.org/10.1029/2002JA009504`.

A. A. van Ballegooijen and P. C. H. Martens. Formation and eruption of solar prominences. *ApJ*, 343: 971–984, August 1989. doi: 10.1086/167766.

H. C. van de Hulst. The electron density of the solar corona. *Bull. Astron. Inst. Neth.*, 11:135, February 1950.

H. van der Vorst. Bi-cgstab: A fast and smoothly converging variant of bi-cg for the solution of nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing*, 13(2):631–644, 1992. doi: 10.1137/0913035. URL `http://epubs.siam.org/doi/abs/10.1137/0913035`.

Wenbin Wang, Jiuhou Lei, Alan G. Burns, Michael Wiltberger, Arthur D. Richmond, Stanley C. Solomon, Timothy L. Killeen, Elsayed R. Talaat, and David N. Anderson. Ionospheric electric field variations during a geomagnetic storm simulated by a coupled magnetosphere ionosphere thermosphere (cmit) model.

Geophysical Research Letters, 35(18):n/a–n/a, 2008. ISSN 1944-8007. doi: 10.1029/2008GL035155. URL `http://dx.doi.org/10.1029/2008GL035155`. L18105.

D. F. Webb and A. J. Hundhausen. Activity associated with the solar origin of coronal mass ejections. *Sol Phys*, 108:383–401, September 1987. doi: 10.1007/BF00214170.

D. R. Weimer. Models of high-latitude electric potentials derived with a least error fit of spherical harmonic coefficients. *Journal of Geophysical Research*, 100:19595–19608, October 1995. doi: 10.1029/95JA01755.

D. R. Weimer. A flexible, IMF dependent model of high-latitude electric potentials having "space weather" applications. *Geophysical Research Letters*, 23:2549–2552, 1996. doi: 10.1029/96GL02255.

D. R. Weimer. An improved model of ionospheric electric potentials including substorm perturbations and application to the Geospace Environment Modeling November 24, 1996, event. *Journal of Geophysical Research*, 106:407–416, January 2001. doi: 10.1029/2000JA000604.

D. R. Weimer. Improved ionospheric electrodynamic models and application to calculating joule heating rates. *Journal of Geophysical Research: Space Physics*, 110(A5):n/a–n/a, 2005. ISSN 2156-2202. doi: 10.1029/2004JA010884. URL `http://dx.doi.org/10.1029/2004JA010884`.

D. R. Weimer, D. M. Ober, N. C. Maynard, M. R. Collier, D. J. McComas, N. F. Ness, C. W. Smith, and J. Watermann. Predicting interplanetary magnetic field (IMF) propagation delay times using the minimum variance technique. *Journal of Geophysical Research (Space Physics)*, 108:1026, January 2003. doi: 10.1029/2002JA009405.

M. Wiltberger, W. Wang, A. G. Burns, S. C. Solomon, J. G. Lyon, and C. C. Goodrich. Initial results from the coupled magnetosphere ionosphere thermosphere model: magnetospheric and ionospheric responses. *Journal of Atmospheric and Solar-Terrestrial Physics*, 66:1411–1423, October 2004. doi: 10.1016/j.jastp.2004.03.026.

M. Wiltberger, L. Qian, C.-L. Huang, W. Wang, R. E. Lopez, A. G. Burns, S. C. Solomon, Y. Deng, and Y. Huang. CMIT study of CR2060 and 2068 comparing L1 and MAS solar wind drivers. *Journal of Atmospheric and Solar-Terrestrial Physics*, 83:39–50, July 2012. doi: 10.1016/j.jastp.2012.01.005.

B. E. Wood, R. A. Howard, A. Thernisien, S. P. Plunkett, and D. G. Socker. Reconstructing the 3D Morphology of the 17 May 2008 CME. *Sol Phys*, 259:163–178, October 2009. doi: 10.1007/s11207-009-9391-0.

S. Yashiro, N. Gopalswamy, G. Michalek, O. C. St. Cyr, S. P. Plunkett, N. B. Rich, and R. A. Howard. A catalog of white light coronal mass ejections observed by the SOHO spacecraft. *Journal of Geophysical Research (Space Physics)*, 109:A07105, July 2004. doi: 10.1029/2003JA010282.

E. Yiğit and A. J. Ridley. Effects of high-latitude thermosphere heating at various scale sizes simulated by a nonhydrostatic global thermosphere-ionosphere model. *Journal of Atmospheric and Solar-Terrestrial Physics*, 73:592–600, April 2011. doi: 10.1016/j.jastp.2010.12.003.

B. Zhang, W. Lotko, M.J. Wiltberger, O.J. Brambles, and P.A. Damiano. A statistical study of magnetosphereionosphere coupling in the lyonfeddermobarry global {MHD} model. *Journal of Atmospheric and Solar-Terrestrial Physics*, 73(56):686 – 702, 2011. ISSN 1364-6826. doi: http://dx.doi.org/10.1016/j.jastp.2010.09.027. URL `http://www.sciencedirect.com/science/article/pii/S1364682610002877`.

S. Zidowitz. The Coronal Structure in August 1996-A Tomographic Reconstruction. In A. Wilson, editor, *Fifth SOHO Workshop: The Corona and Solar Wind Near Minimum Activity*, volume 404 of *ESA Special Publication*, page 757, 1997.

S. Zidowitz. Coronal structure of the Whole Sun Month: A tomographic reconstruction. *Journal of Geophysical Research*, 104:9727–9734, May 1999. doi: 10.1029/1998JA900099.

S. Zidowitz, B. Inhester, and A. Epple. Tomographic inversion of coronagraph images. In D. Winterhalter, J. T. Gosling, S. R. Habbal, W. S. Kurth, and M. Neugebauer, editors, *American Institute of Physics Conference Series*, volume 382 of *American Institute of Physics Conference Series*, pages 165–168, July 1996. doi: 10.1063/1.51466.