



**ON THE USE OF MULTI-MODEL ENSEMBLE
TECHNIQUES FOR IONOSPHERIC
AND THERMOSPHERIC CHARACTERISATION**

by

SEAN ELVIDGE, MSci.

A thesis submitted to
The University of Birmingham

For the Degree
DOCTOR OF PHILOSOPHY

Space Environment and Radio Engineering Group
Electronic, Electrical and Systems Engineering
College of Engineering and Physical Sciences,
University of Birmingham
August 2014

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

ABSTRACT

Space weather can have a negative impact on a number of radio frequency (RF) systems, with mitigation by ionospheric and thermospheric modelling one approach to improving system performance. However, before a model can be adopted operationally its performance must be quantified. Taylor diagrams, which show a model's standard deviation and correlation, have been extended to further illustrate the model's bias, standard deviation of error and mean square error in comparison to observational data. By normalising the statistics, multiple parameters can be shown simultaneously for a number of models. Using these modified Taylor diagrams, the first known long term (one month) comparison of three model types – empirical, physics and data assimilation - has been performed. The data assimilation models performed best, offering a statistically significant improvement in performance. One physics model performed sufficiently well that it is a viable background model option in future data assimilation schemes. Finally, multi-model thermospheric ensembles (MMEs) have been constructed from which the thermospheric forecasts exhibited a reduced root mean square error compared to non-ensemble approaches. Using an equally weighted MME the reduction was 55% and using a mean square error weighted approach the reduction was 48%.

To,

Jed, Leo, Josh, C.J., Toby,
Charlie, Donna, Sam & Will.

ACKNOWLEDGEMENTS

The work described in this thesis would not have been possible without the help of a wide range of people. My parents have always been a massive support to me, and have done anything and everything to help me continue my education. My wife, Katie, is a constant support, and without her I would surely be lost. I have always felt tremendously fortunate to have had incredible teachers throughout my life, from Primary School to University. I would like to thank you all for your hard work, kindness and patience.

My PhD was undertaken in the Space Environment and Radio Engineering (SERENE) group at the University of Birmingham. It has been a fantastic experience being there, better than I could have ever imagined. A special thank-you to Paul Cannon OBE, our former director, Costas Constantinou and ‘like’ Geoff de Villiers. During my thesis I was lucky enough to go to the Los Alamos National Laboratory Space Weather Summer School. It was there that, working with Humberto Godinez, I first developed the multi-model ensemble work. I am hugely indebted to Humberto for his (continuing) help and I look forward to our future collaboration. A thank-you also goes to Chris Mannix, who is always available for me to bounce ideas off and drink coffee with. I hope I can return the favour to you as you write your thesis.

Finally, to my PhD supervisor, Matthew Angling. I am tremendously grateful for the amount of help and support you have provided over the last three years. I’ve always found discussions with you to be enlightening and you have taught me more about the English language than anyone.

TABLE OF CONTENTS

Abstract.....	i
Acknowledgements	iii
Table of Contents	iv
List of Figures.....	vi
List of Tables.....	xii
List of Abbreviations.....	xiii
1. Introduction	1
2. The Ionosphere	3
2.1. Introduction	3
2.2. Ionospheric Nomenclature.....	4
2.3. Ionospheric Structure.....	4
2.3.1. D Region.....	6
2.3.2. E Region	7
2.3.3. F Region	7
2.4. Storm Time Ionosphere	8
2.5. The Plasmasphere	10
2.6. Refractive Index	10
2.7. Data Types.....	12
2.7.1. Global Navigation Satellite System Data	12
2.7.2. Ionosondes	15
3. Ionospheric Models	18
3.1. Empirical Models	18
3.1.1. International Reference Ionosphere (IRI).....	19
3.1.2. NeQuick.....	20
3.1.3. NRLMSISE-00	21
3.2. Physics-Based Models.....	22
3.2.1. TIE-GCM	23
3.2.2. GITM.....	23
3.3. Data Assimilation	24
3.3.1. Least Squares	26
3.3.2. Weighted Least Squares	27
3.3.3. Kalman Filters	28
3.3.4. Electron Density Assimilative Model (EDAM)	31
3.3.5. GPS Ionospheric Inversion (GPSII)	32
3.3.6. NeQuick Data Ingestion Model.....	32
3.4. Multi-Model Ensemble Introduction.....	33

3.5.	Conclusions	34
4.	How To Compare Ionospheric Models.....	35
4.1.	Current Model Comparison Techniques.....	35
4.2.	Taylor Diagrams	37
4.3.	Modified Taylor Diagrams	40
4.4.	Conclusions	41
5.	Model Comparisons.....	43
5.1.	Testing Scenario	45
5.2.	Filtering Observations and Model Results	48
5.3.	Observations	50
5.4.	Comparison of IRI, EDAM, GPSII and mNeQuick.....	55
5.4.1.	Accuracy of Model Values at Hermanus.....	55
5.4.2.	Accuracy of Model Values at Louisvale	66
5.4.3.	Accuracy of Model Values at Madimbo.....	68
5.5.	Comparisons Including TIE-GCM and GITM	70
5.6.	Comparisons Including Model Combinations.....	73
5.7.	Conclusions	77
6.	Multi-Model Ensembles	79
6.1.	Atmospheric Neutral Densities.....	79
6.2.	Test Scenario	80
6.3.	Models and Observations	81
6.3.1.	CHAMP	83
6.4.	Model Comparison	84
6.5.	Rationale of MMEs	87
6.6.	Constructing an MME	90
6.6.1.	Equally Weighted MME.....	90
6.6.2.	Weighted MME	90
6.7.	Using the MME for Improved Atmospheric Density Modelling	91
6.7.1.	MME Construction.....	91
6.7.2.	Using the MME as the Initial Conditions for TIE-GCM	98
6.8.	Conclusions	106
7.	Conclusions and Future work	108
Appendix A.	The relationship between electron density and plasma frequency	112
Appendix B.	Estimation of TEC from GNSS Measurements.....	114
List of References	123

LIST OF FIGURES

Figure 1. A representation of the ionosphere showing the four regions, and how they vary diurnally and with respect to the solar cycle [Angling et al., 2013].....	6
Figure 2. Distribution of GNSS receivers available from IGS [IGS Central Bureau, 2010]. ..	13
Figure 3. A sample ionogram [Lowell Digisonde International, 2014a]. Ionosonde sounding frequency is plotted on the x-axis and virtual height on the y-axis. Pixel intensity shows the signal amplitude and Doppler shift is plotted as the colour shade. The red-yellow-white scale shows the O wave polarization whilst the blue-green-grey scale shows the X wave.	16
Figure 4. Existing and planned ionosonde stations by Digisonde [Lowell Digisonde International, 2014b].	17
Figure 5. Observed (black curve) and modelled (colour curves) of hmF2 at a particular testing site. Reproduced from Shim et al. [2011].....	36
Figure 6. Example Taylor diagram for models ‘A’ through ‘H’. The dashed line (labelled observed), shows the standard deviation of the observations. The dotted line semi-circles, originating from the intersection of the observed standard deviation (dashed line) and the horizontal axis, show the centred pattern RMS value. Considering a specific model, the radial distance from the origin shows the models’ standard deviations and the azimuthal angle corresponds to the correlation between the model and the observation.	39
Figure 7. Example modified Taylor diagram.	42
Figure 8. Locations of the four Republic of South Africa ionosondes. Louisvale and Hermanus are ~700 km from Grahamstown. Madimbo is ~1300 km from Grahamstown.	45

Figure 9. Location of the Sub-Saharan Africa GNSS stations used in this test.	47
Figure 10. Time series plot of the sunspot number (red) and Ap (blue) for September 2011. Ap values above 29 exceed the storm threshold as defined by the Space Weather Prediction Center [2009].	48
Figure 11. Reported foF2 time series from the Grahamstown ionosonde in the RSA, subject to the filtering described in Section 5.2.	51
Figure 12. Reported hmF2 time series from the Grahamstown ionosonde in the RSA, subject to the filtering described in Section 5.2.	51
Figure 13. Black line is the observed F10.7 values over September 2011, the red line is a 3 rd order polynomial fitted to the data.	52
Figure 14. Modified Taylor diagram to show the model bias and standard deviation of errors. The azimuthal angle represents correlation, the radial distance the standard deviation and the semicircles centred at the ‘Observed’ marker the standard deviation of the errors. The colour scale shows the bias (mean of model minus mean of truth). To plot multiple parameters each quantity is normalized, the original values can be reformed using the corresponding ‘factors’ in the top right of the diagram. The diagram shows the second-order statistics for IRI, EDAM, GPSII and mNeQuick at Hermanus for the parameters foF2, hmF2 and h(0.8foF2).....	58
Figure 15. Modified Taylor diagram for Hermanus with second-order statistics split between day (08UT - 15UT), night (22UT - 5UT) and dawn/dusk (03UT – 08UT and 15UT – 20UT) for EDAM. The details of how to read the diagram are described in Figure 14.	60

Figure 16. Modified Taylor diagram of the second order statistics for EDAM at Hermanus. Each day of the test study (September 2011) is plotted individually. The details of how to read the diagram are described in Figure 14. 62

Figure 17. Modified Taylor diagram of the second order statistics for GPSII at Hermanus. Each day of the test study (September 2011) is plotted individually. The over plotted ovals are the same as in Figure 16. The details of how to read the diagram are described in Figure 14. 63

Figure 18. Modified Taylor diagram of the second order statistics for mNeQuick at Hermanus. Each day of the test study (September 2011) is plotted individually. The details of how to read the diagram are described in Figure 14. 64

Figure 19. Modified Taylor diagram of the second order statistics for IRI at Hermanus. Each day of the test study (September 2011) is plotted individually. The details of how to read the diagram are described in Figure 14. 65

Figure 20. Modified Taylor diagram for IRI, EDAM, GPSII and mNeQuick at Louisvale for the parameters foF2, hmF2 and h(0.8foF2). The details of how to read the diagram are described in Figure 14..... 68

Figure 21. Modified Taylor diagram for IRI, EDAM, GPSII and mNeQuick at Madimbo for the parameters foF2, hmF2 and h(0.8foF2). The details of how to read the diagram are described in Figure 14..... 70

Figure 22. Modified Taylor diagram for IRI, EDAM, GPSII, mNeQuick, GITM and TIE-GCM at Hermanus for the parameters foF2, hmF2 and h(0.8foF2). The details of how to read the diagram are described in Figure 14. 72

Figure 23. Time series of GITM’s foF2 values across the month of September 2011. The nighttime values have unphysical characteristics..... 73

Figure 24. Modified Taylor diagram for IRI, EDAM, GPSII, mNeQuick and the average combination of their parameter time series at Hermanus for the parameters foF2, hmF2 and h(0.8foF2). The details of how to read the diagram are described in Figure 14. 74

Figure 25. Example electron density height profile, black solid line is from EDAM and dashed line from GPSII. The red solid line is the average of the two. 76

Figure 26. Modified Taylor diagram for EDAM, GPSII and the average combination of their height profiles at Hermanus for the parameters foF2, hmF2 and h(0.8foF2). The details of how to read the diagram are described in Figure 14. 76

Figure 27. Ap and F10.7 for the neutral density MME test study, August 28th – September 1st 2009. The spike in Ap is due to a geomagnetic storm..... 81

Figure 28. Neutral densities, derived from CHAMP accelerometer data for August 28th – September 1st 2009. The time step of the x-axis is 30 minutes, which was approximately one third of the CHAMP orbital period. This is the cause of the apparent oscillation in the neutral densities. 84

Figure 29. Modified Taylor diagram for NRLMSISE-00 (MSIS), TIE-GCM and GITM for neutral density, compared with CHAMP. The details of how to read the diagram are described in Figure 14..... 85

Figure 30. Plot showing the neutral density observation data from CHAMP (black) with the three model outputs, NRLMSISE-00 (orange), GITM (red) and TIE-GCM (blue) for the study period. 86

Figure 31. Idealized visualization of the basic MME scenarios, based on the cartoon in Hagedorn et al. [2005]. Two models are shown (red and black lines) for when 1) the MME would produce the best estimate, 2) one model provides the best estimate and 3) no model provides a good estimate. 89

Figure 32. Modified Taylor diagram of NRLMSISE-00 (MSIS), TIE-GCM and GITM for the neutral density test. The model time series have been restricted to times when the Ap was less than or equal to 3. These time series are used to generate the weighted MME. 92

Figure 33. The CHAMP derived neutral densities are shown in black. In green are the neutral density values found from the average MME. 95

Figure 34. The CHAMP derived neutral densities are shown in black. In purple are the neutral density values found from the weighted MME. 96

Figure 35. Modified Taylor diagram for NRLMSISE-00 (MSIS), TIE-GCM, GITM and TIE-GCM started with an MME for neutral density, compared with CHAMP. The details of how to read the diagram are described in Figure 14..... 97

Figure 36. Flow chart of the procedure for running TIE-GCM using the MME as its initial conditions for a six hour forecast. 100

Figure 37. Procedure for finding the TIE-GCM forecast using an MME as its initial conditions for the August 28th – September 1st 2009 neutral density study. The “run TIE-GCM MME forecast” process refers to the procedure described in Figure 36..... 101

Figure 38. Modified Taylor diagram for NRLMSISE-00 (MSIS), TIE-GCM, GITM and for TIE-GCM using the MME (both average and weighted) for its initial conditions every six hours, compared with the CHAMP observations. The details of how to read the diagram are described in Figure 14..... 104

Figure 39. Top panel shows the neutral density from the CHAMP observations, from the original TIE-GCM run and the equal combination MME. The bottom panel shows the CHAMP observations and the new TIE-GCM output, using the MME as the initial condition every six hours. 105

Figure 40. Top panel shows the neutral density from the CHAMP observations, from the original TIE-GCM run and the weighted MME. The bottom panel shows the CHAMP observations and the new TIE-GCM output, using the MME as the initial condition every six hours. 106

LIST OF TABLES

Table 1. Locations of ionosonde data assimilated in the test.	46
Table 2. GNSS stations assimilated in the test.	46
Table 3. Number of valid data points produced by the ionosondes in the RSA for September 2011 subject to the filtering described in Section 5.2.....	50
Table 4. Results from the ADF test for the model time series. For each foF2 time series, the value is greater than the critical value for the 95% confidence level (-3.96). However, after detrending, the time series is stationary.....	54
Table 5. Z values from comparing the correlation coefficients for foF2 for each of the four models. Values greater than or equal to 1.96 are significant at the 95% level.	56
Table 6. Suggested correlation lengths for use with global models. Reproduced from Table 3 in McNamara [2009].....	66
Table 7. NRLMSISE-00, TIE-GCM and GITM model outputs. mmr is the mass mixing ratio.	82
Table 8. NRLMSISE-00, TIE-GCM and GITM model (solar proxy) inputs.....	83
Table 9. Model skill and associated weighting (calculated by the inverse of model skill, Equation (6.2)) for use in the weighted MME.	93

LIST OF ABBREVIATIONS

ADF	Augmented Dickey-Fuller
Ap	Index for severity of the magnetic disturbances in near-Earth space
BLUE	Best linear unbiased estimate
CCIR	Consultative Committee on International Radio
CHAMP	CHALLENGING Minisatellite Payload
COSPAR	Committee on Space Research
EDAM	Electron Density Assimilative Model
EUV	Extreme ultraviolet
F10.7	Solar flux at a wavelength of 10.7 cm (2800 MHz) at the Earth's orbit
GITM	Global Ionosphere-Thermosphere Model
GLONASS	Globalnaya Navigatsionnaya Sputnikovaya Sistema
GNSS	Global navigation satellite system
GPS	Global Positioning System
GPSII	GPS Ionospheric Inversion
HF	High frequency (3 – 30 MHz)
HPI	Hemispheric power index
ICTP	The Abdus Salam International Centre for Theoretical Physics
IGAM	Institute for Geophysics, Astrophysics and Meteorology
IGS	International GNSS Service
IRI	International Reference Ionosphere
ITU-R	International Telecommunication Union Radiocommunication Sector
Kp	Proxy for geomagnetic activity, similar to Ap.
LEO	Low Earth orbit
M(3000)F2	Ratio of maximum usable frequency at 3000 km to foF2
MME	Multi-model ensemble
mNeQuick	Data ingestion version of NeQuick
MSE	Mean square error

MSIS	Mass Spectrometer and Incoherent Scatter radar
NCAR	National Center for Atmospheric Research
NRLMSISE-00	Naval Research Laboratory Mass Spectrometer Incoherent Scatter Exosphere - 2000
OI	Optimal interpolation
OTHR	Over-the-horizon radar
REA	Reliability ensemble averaging
RF	Radio frequency
RMS	Root mean square
RO	Radio occultation
SID	Sudden ionospheric disturbance
TEC	Total electron content
TECU	Total electron content unit
TID	Travelling ionospheric disturbance
TIE-GCM	Thermosphere Ionosphere Electrodynamics – Global Circulation Model
URSI	International Union of Radio Science
WLS	Weighted least squares

1. INTRODUCTION

None but those who have experienced them can conceive of the enticements of science... In a scientific pursuit there is continual food for discovery and wonder.

- Mary Wollstonecraft Shelley, *Frankenstein*

Space weather refers to “conditions on the Sun and in the solar wind, magnetosphere, ionosphere, and thermosphere that can influence the performance and reliability of space-borne and ground-based technological systems and can endanger human life or health. Adverse conditions in the space environment can cause disruption of satellite operations, communications, navigation, and electric power distribution grids, leading to a variety of socioeconomic losses.” [National Space Weather Program, 1995]. Variations in solar and geomagnetic activity can cause ionospheric variability and radio frequency (RF) systems that operate via or through the ionosphere at frequencies of ~2 GHz and below can be severely affected. Thus space weather poses a threat to systems such as global navigation satellite systems (GNSS), high frequency (HF) communications, space-based Earth observation radars and space situational awareness radars. Therefore, comprehensive, global and timely specifications of the Earth's ionosphere are required to ensure the effective operation, planning and management of these systems. Many techniques have been developed to measure the ionosphere; including ground and space-based ionosondes [Titheridge, 1988] and the use of global navigation satellite system (GNSS) measurements made with both ground and space-based receivers [Dow et al., 2005].

Ionospheric models can be used to help mitigate against these effects. There are a wide range of ionospheric models in active use and development utilizing a number of modelling

techniques. The models can be split into three main categories: empirical, physics-based (i.e. solving equations governing the underlying physical system) and data assimilation (an optimal combination of one of the first two approaches with ionospheric data). Empirical models are widely used due to their relative simplicity and good median performance [e.g. *Llewellyn and Bent, 1973; Bilitza et al., 1988*]. However they are unsuitable for nowcasting or medium term (hours to days) forecasting since only very basic data is considered in the modelling procedure. In general they do not respond to varying space weather conditions (such as the onset of a geomagnetic storm). Physics-based models strive to solve the equations which underpin the physical system of the ionosphere [e.g. *Richmond et al., 1992; Ridley et al., 2006*]. Such models also rely on initial and boundary conditions from the thermosphere, plasmasphere and magnetosphere, as well as proxies for solar activity, to drive the models. Errors in these measurements can introduce errors into the model output. Data assimilation models attempt to optimally combine individual measurements and a background model of the ionosphere [e.g. *Angling and Khattatov, 2006; McNamara et al., 2013*]. They rely on the availability of data (such as from GNSS stations and ionosondes) and sophisticated statistical techniques to combine the data and background model.

To gain insight into the skill of the various ionospheric modelling types and techniques, suitable metric(s) are required. The ‘modified Taylor diagram’ [*Taylor, 2001; Elvidge et al., 2014*], a visualization technique for analysing a variety of parameters from multiple models simultaneously, is described in this work. Using modified Taylor diagrams, the testing results from a variety of different models are presented [*McNamara et al., 2013; Elvidge and Angling, 2014*]. Finally, multi-model ensembles (MMEs) are described and used, to show how the combination of multiple models can provide improved results [*Elvidge et al., 2013*].

2. THE IONOSPHERE

The beauty of the natural world lies in the details.

- Natalie Angier, *The Beauty of the Beastly*

2.1. Introduction

From approximately 60 km above the Earth's surface to 1500 km [Angling *et al.*, 2013], atoms and molecules in the atmosphere are ionised by radiation from the Sun. The process of ionisation removes electrons from these atoms and molecules, leaving positively charged ions and free electrons in a region of the atmosphere called the ionosphere. The ionosphere is thus an ionised gas, i.e. a plasma. Since the ionisation is caused by the Sun, solar activity and ionospheric variability are closely linked [Davies, 1990].

The number of free electrons in a volume, the electron density, depends on the speed of the electron production and loss processes [Hunsucker and Hargreaves, 2003]. This can be expressed by the following equation:

$$\frac{\partial N_e}{\partial t} = q - L - \nabla \cdot (N_e \mathbf{v}). \quad (2.1)$$

Where N_e is the electron density, q the production rate, L the loss rate and $\nabla \cdot (N_e \mathbf{v})$ the loss rate of electrons due to transport (\mathbf{v} the electron mean velocity). Chapman [1931] developed a simple mathematical model for how ionospheric layers, or regions, form based on the zenith angle of the Sun. The Chapman model uses the assumption that ionizing radiation can be characterized by a single frequency. It also depends upon there being a balance between the formation of free electrons and their loss by recombination (the process by which electrons and

ions join together) and that the atmosphere is made up of a single species. Although a simple model, it describes the main characteristics of the different regions. The approach is also useful for describing the shape of the vertical profile of electron densities in the ionosphere.

2.2. Ionospheric Nomenclature

When discussing the ionosphere there are a number of abbreviations which are commonly used to describe various physical parameters. The main abbreviations are based around the prefixes: ‘Nm’ (peak electron density), ‘fo’ (ordinary mode (O mode) critical frequency, see Section 2.6 for further details), ‘fx’ (extraordinary mode (X mode) critical frequency) and ‘hm’(height), with suffixes referring to the ionospheric regions: E, F1 and F2.

For example, NmF2 is the peak electron density in the F2 region, foE is the O mode critical frequency in the E region and hmF1 is the height of the peak electron density/plasma frequency in the F1 region. This naming convention derives from the ionosonde community [*Piggott and Rawer, 1978*] and there are many more names than just the ones mentioned here.

2.3. Ionospheric Structure

Extreme ultraviolet (EUV) and X-ray radiation from the Sun ionises atoms and molecules in the atmosphere, releasing free electrons. Since the density of the atmosphere at high altitudes is relatively low, a large number of electrons remain free. It is because of this that the ionosphere forms. Increases in the amount of solar radiation, such as during a solar storm, will affect the quantity of electron/ion pairs produced.

The behaviour of the ionosphere is dependent on processes in the thermosphere and the Earth's magnetic field [Mitchell, 2013]. With increasing altitude, the thermospheric effects decrease, and the magnetic field effects become dominant. The direction of the Earth's magnetic field, with respect to the ionosphere, is important in understanding how different geographic regions are affected. At low latitudes, the magnetic field is almost horizontal, whereas at high latitudes it is vertical. At extremely high latitudes the magnetic field is open to the interplanetary magnetic field (IMF), and the full effects of the Sun can be seen. It is because of this that aurora generally occur at high latitudes.

The movement of the ionosphere is again dependent upon geographic location. For example, near the equator when the zonal electric field is eastwards it causes plasma to move upwards. Whereas at high latitudes, the electric field generated by the IMF, $(E \times B)/B^2$, controls the horizontal speed of the ionosphere. A more detailed explanation of ionospheric structure and the physics of the ionosphere can be found in *Hargreaves* [1979].

The vertical profile of the ionosphere varies both temporally and spatially. It is traditionally split into four 'regions' or 'layers', D, E, F1 and F2 (Figure 1). The characteristics of each region is also dependent upon geomagnetic conditions. Each of the regions are described in more detail in the following sections.

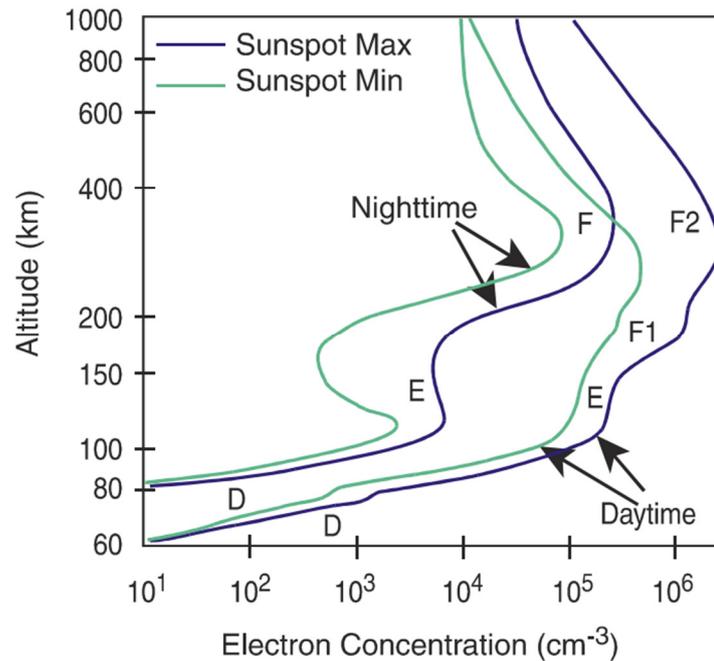


Figure 1. A representation of the ionosphere showing the four regions, and how they vary diurnally and with respect to the solar cycle [Angling et al., 2013].

2.3.1. D Region

The D region is the lowest ionospheric region, typically occurring between ~60 and ~90 km. The region is produced almost entirely by direct ionisation of *NO* [Mitchell, 2013]. It can be seen in Figure 1 that it does not have a distinct peak. During the day the region has a peak value of around 2×10^8 electrons/m³. It disappears at night since the high neutral density results in a high recombination rate of the electron/ion pairs.

Increases in solar X-rays (such as during a solar flare), increase D region absorption. Absorption is the process by which the energy from radio waves is converted into heat and electromagnetic noise [Davies, 1990]. Such increases can potentially lead to sudden ionospheric disturbances

(SIDs). The region does not support HF propagation due to the low plasma frequency (see Section 2.6).

2.3.2. E Region

The E region, from ~90 to ~130 km has a peak electron density (NmE) of 1×10^{11} electrons/m³. Soft X-rays (wavelengths between 0.8 nm and 1.4 nm) and ultraviolet radiation (100 nm to 150 nm) are the main form of ionising radiation of O_2 and N_2 in this region [Schunk, 1996]. The region is solar zenith angle dependent, and as such largely disappears at night. It is the region best modelled by a Chapman layer and few modern techniques focus on the region [Solomon, 2006].

2.3.3. F Region

The F region is the densest ionospheric layer and the one which has most impact on radio systems. In the day the region splits into two ‘sub’ regions, F1 and F2 (described below), whereas at night it collapses to just the single F region. It is formed by O and N_2 ionised by EUV radiation, with wavelengths between 10 nm and 90 nm [Parsons, 2006].

The F1 region extends from ~130 km to ~210 km, and typically has a peak electron density (NmF1) of approximately 2×10^{11} electrons/m³. Even though the F1 layer behaves approximately as a Chapman layer, its peak values are affected by high solar activity [Rastogi, 1958].

Extending from ~210 km to ~600 km (the ‘upper bound’ of the ionosphere is not well defined and various texts place it between 1000 km and 2000 km [Yau *et al.*, 1996; Angling *et al.*, 2013; Mitchell, 2013]), the F2 region generally exhibits the greatest electron density. The NmF2 is approximately 1×10^{12} electrons/m³ in the daytime and drops to 5×10^{10} electrons/m³ at night. The region persists through the night due to the lower collision rate (thus the longer lifetime of electron/ion pairs) and transport of electrons from other regions.

The F2 region contains many anomalous features when compared to Chapman layers and can be particularly difficult to model (in contrast to the E region). These include the equatorial (Appleton) anomaly, which is a region of heightened ionospheric density at $\pm 20^\circ$ of the magnetic equator, and the winter anomaly where the F2 ionisation is lower in the summer than the winter. These anomalies, and others, are described more thoroughly in Davies [1990], which should also be consulted for a more detailed description of the ionosphere and its processes.

2.4. Storm Time Ionosphere

Changes in solar activity can have a dramatic impact on electron densities in the ionosphere, as well as neutral/ion densities and temperature [Gorney, 1990]. During a geomagnetic storm the Earth’s magnetosphere is compressed by the solar wind and electric fields flow along the magnetic field lines into the ionosphere [Buonsanto, 1999]. High energy particles also precipitate into the ionosphere and thermosphere. The ionosphere responds to these affects in a variety of ways, which vary spatially. Hunsucker and Hargreaves [2003] describe the three main stages of an ionospheric storm on electron density across multiple latitudes as:

1. Initial phase – electron density and content increase.

2. Main phase – electron density and content decrease below normal values.
3. Recovery phase – electron density and content gradually return to normal.

In the initial phase increased ionisation, particularly on the dayside D region of the Earth, causes SIDs which affect HF signals, degrading radio communication [Belrose and Thomas, 1968]. The F region electron density sees a sudden increase due to a decrease in the recombination rate which occurs because of a reduction in O₂ and N₂. In the main phase the electron density drops due to an increase in the recombination rate because of an increase of O₂ and N₂. These changes in the recombination rate are caused by changes in the altitude of O₂ and N₂ which control the dominant chemical reactions. The ionosphere then slowly returns to normal during the recovery phase. The speed of recovery is slow due to the time-scales of molecular diffusion in the atmosphere [Fuller-Rowell *et al.*, 2001]. These effects occur globally, however, there are other effects specific to particular ionospheric regions.

In low-latitudes rapid changes in the electric field, inside the magnetosphere, penetrate into the F region. This causes sudden downwelling or upwelling which creates instabilities called spread F. Also at low latitudes, the dayside eastward and nightside westward electric currents meet which can force ionospheric plasma upwards [Knipp, 2011].

In high latitudes, E region and F region ionospheric effects are most prominent. Electric fields mapped along magnetic field lines transport plasma on the dayside to the nightside, across the polar cap. This can cause large trails of ionization, which can be further enhanced by particle precipitation which is guided by the Earth's magnetic field lines into the auroral regions. During the initial phase of a geomagnetic storm, the polar cap potential increases which can lead to an

expansion of the auroral zone [Kavanagh *et al.*, 2004]. Finally, solar proton events (SPEs), which are mainly MeV energised protons, ionize low altitude regions of the polar caps. These events are named polar cap absorption (PCA) events.

There are a number of reviews of solar storm effects on the ionosphere and thermosphere including: Fuller-Rowell *et al.* [1994], Lastovicka [1996], Buonsanto [1999], Hunsucker and Hargreaves [2003] and Knipp [2011]. These works should be consulted for more details on the processes described in this section as well as other, undescribed, phenomena.

2.5. The Plasmasphere

The plasmasphere is the region above the ionosphere. Its lower boundary is defined as the region where protons replace oxygen as the majority species in the atmosphere. This occurs between 1000 and 2000 km. It extends to between four and six Earth radii (19000 to 32000 km) to the plasmopause. There is very little ion production in this region but ions are produced in the ionosphere during the day, and diffuse upwards [Davies, 1990]. Modelling of the plasmasphere is particularly important to physics-based models as it is a boundary condition for the ionosphere.

2.6. Refractive Index

How an electromagnetic wave propagates through a medium is dependent upon the refractive index of the medium. The ionosphere is a cold, magnetized plasma and how a wave propagates through it is described by the Appleton-Hartree-Lassen equation. The Appleton-Hartree-Lassen equation (sometimes referred to just as the Appleton [Ratcliffe, 1959], Appleton-Hartree

[Budden, 1985], or Appleton-Lassen [Rawer and Suchy, 1976] equation) describes the refractive index for electromagnetic wave propagation [Lassen, 1927; Hartree, 1929; Appleton, 1932; Davies, 1990]. The refractive index (n) can be calculated using the simplified form of the Appleton-Hartree-Lassen equation [Angling *et al.*, 2013]:

$$n^2 = 1 - \frac{2X(1 - X)}{2(1 - X) - Y_T^2 \pm \sqrt{Y_T^4 + 4(1 - X)^2 Y_L^2}}. \quad (2.2)$$

Where

$$X = \frac{N_e q_e^2}{\epsilon_0 m \omega^2}, \quad (2.3)$$

$$Y_T = \frac{q_e B_T}{m \omega}, \quad (2.4)$$

$$Y_L = \frac{q_e B_L}{m \omega}. \quad (2.5)$$

and N_e is the electron density, q_e is the charge of the electron (1.6×10^{-19} coulomb), ϵ_0 is the vacuum permittivity (8.85×10^{-12} farad-meter⁻¹), m is the electron mass (9.11×10^{-31} kg), ω is the angular frequency, B_L is the longitudinal component of the magnetic field and B_T is the transverse component of the magnetic field. When a radio wave begins to enter the ionosphere the Earth's magnetic field causes it to split into two characteristic components [Angling *et al.*, 2013]. The two rays are referred to as ordinary (O) and extraordinary (X) waves.

The positive part of the plus/minus sign in the denominator of Equation (2.2) refers to the O wave whilst the negative part refers to the X wave. If one assumes no magnetic field then a relationship between electron density and plasma frequency (f_p) is (derived in Appendix A):

$$f_p = 9\sqrt{N_e}. \quad (2.6)$$

Thus the frequency at which a wave is reflected or refracted by the ionosphere is dependent upon the electron density. Radio waves with frequencies greater than the maximum plasma frequency in the ionosphere will be refracted, but nonetheless, pass through the ionosphere (such as for communication with satellites). Waves with frequencies less than the maximum plasma frequency will be reflected by the ionosphere (such as for over-the-horizon radar (OTHR)). Different regions of the ionosphere have local maxima in electron density and thus waves can be reflected from different heights of the ionosphere.

2.7. Data Types

A wide range of measurement techniques have been developed for measuring the ionosphere. Examples include the use of GNSS [*Hofmann-Wellenhof et al.*, 2001], ionosondes [*Reinisch*, 1996] and incoherent scatter radars [*Rishbeth and Williams*, 1985].

2.7.1. Global Navigation Satellite System Data

A GNSS is a global coverage satellite navigation system. Currently there are only two fully operational GNSS systems: the USA's NAVSTAR Global Positioning System (GPS) and the Russian Globalnaya Navigatsionnaya Sputnikovaya Sistema (GLONASS). However there are two in active deployment; China's BeiDou-2 navigation system (also known as COMPASS) and the European Union's GALILEO system. For global coverage, a satellite navigation system generally requires a satellite constellation of 20 to 30 satellites. GNSS data products are

provided by a number of services including the International GNSS Service (IGS) [Dow *et al.*, 2009] and EUREF [Bruyninx *et al.*, 2001].

There are a large number of GNSS receivers whose data is available for ionospheric modelling. The distribution of receivers from IGS is shown in Figure 2. Despite the large number of GNSS receivers, Saharan Africa has no available receivers through IGS, and very few in the Pacific. This results in a lack of data in those regions.

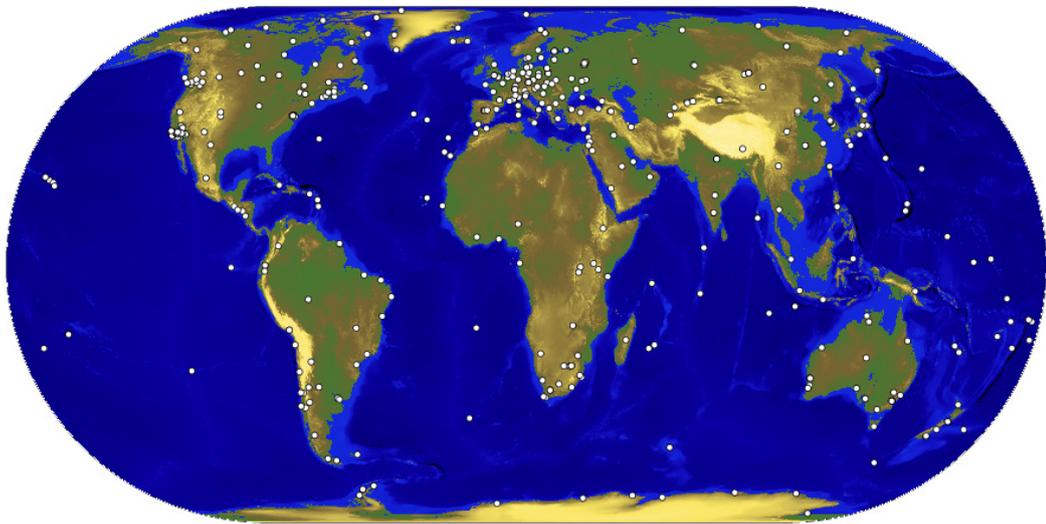


Figure 2. Distribution of GNSS receivers available from IGS [IGS Central Bureau, 2010].

GNSS signals can be used to determine the total electron content (TEC) between the satellite and a receiver on the ground (or in space for radio occultation (RO)). TEC is the line integral of electron density on a path between two points (i.e. between the satellite and the receiver):

$$TEC = \int N_e dl, \quad (2.7)$$

where N_e is electron density. TEC is often measured by the TEC unit (TECU), which is 10^{16} electrons/m².

A dual frequency (L1 (1575.42 MHz) and L2 (1575.42 MHz)) GNSS receiver can record pseudorange (the measured distance between satellite and receiver, assuming the velocity of light) and phase for both frequencies. The phase and pseudorange can be used to derive TEC from GNSS measurements. The phase TEC (Equation (2.8)) gives a 2π ambiguous result (represented by B). Whereas the pseudorange TEC (Equation (2.9)) is a noisy measurement. Therefore, generally the TEC is calculated by using the pseudorange TEC to level the phase TEC (Equation (2.10)).

$$TEC_{phase} = \frac{f_1^2 f_2^2 (\phi_2 - \phi_1)}{40.3(f_2^2 - f_1^2)} + B, \quad (2.8)$$

$$TEC_{pseudorange} = \frac{f_1^2 f_2^2 (P_1 - P_2)}{40.3(f_2^2 - f_1^2)}, \quad (2.9)$$

$$TEC = TEC_{phase} - \overline{TEC_{phase}} + \overline{TEC_{pseudorange}}. \quad (2.10)$$

Where f_i for $i = 1,2$ is the frequency of signal Li (in MHz), ϕ_i is the phase (expressed in metres), B is the unknown ambiguity, P_i is the pseudorange (in metres) and \bar{x} is the mean of x . The derivation of Equations (2.8), (2.9) and (2.10) are shown in Appendix B.

2.7.2. Ionosondes

Ionosondes are radars which are used to examine the ionosphere. Ionosondes typically have four main components [*Lowell Digisonde International*, 2014a]:

- A HF transmitter with frequency coverage from approximately 1 to 40 MHz,
- A HF receiver (which can track the frequency of the transmitter),
- An antenna,
- A computer for data storage and analysis.

An ionosonde produces an ionogram, which is a graph of virtual height (the apparent height of an ionospheric reflections calculated from the time of flight of the radio waves, assuming the wave travels at the speed of light) against frequency.

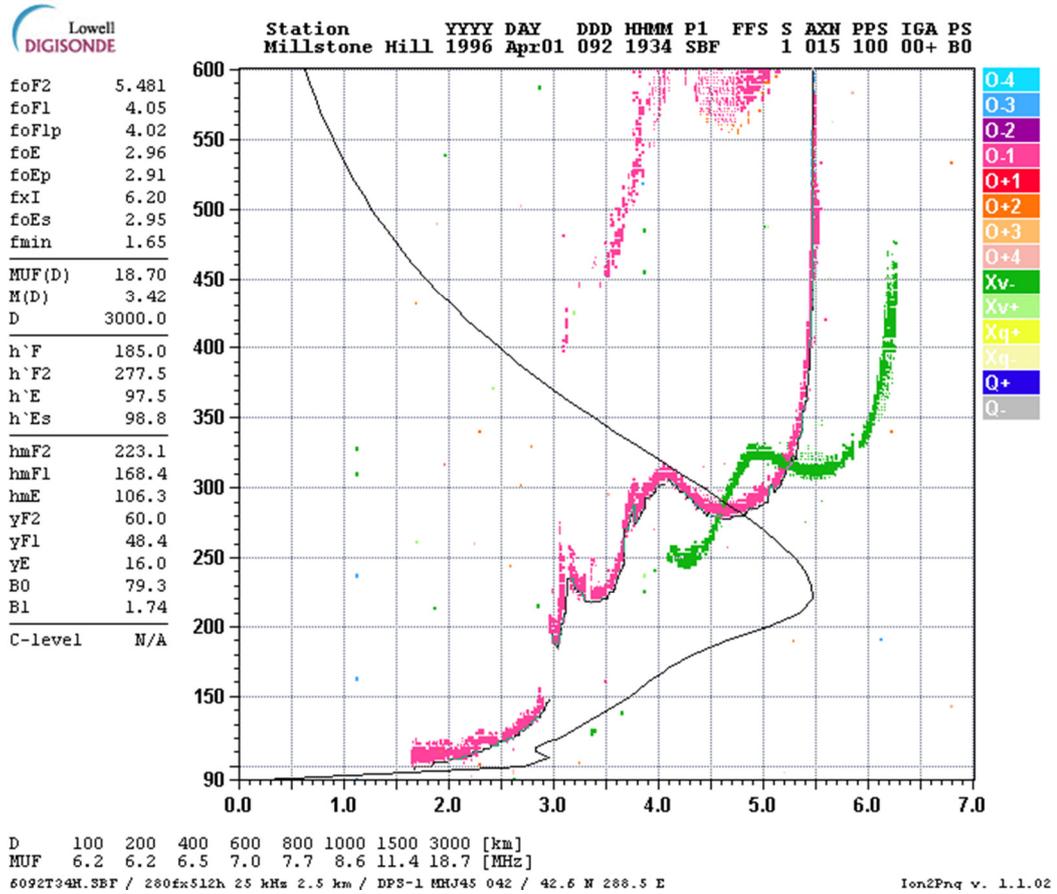


Figure 3. A sample ionogram [Lowell Digisonde International, 2014a]. Ionosonde sounding frequency is plotted on the x-axis and virtual height on the y-axis. Pixel intensity shows the signal amplitude and Doppler shift is plotted as the colour shade. The red-yellow-white scale shows the O wave polarization whilst the blue-green-grey scale shows the X wave.

Both the O and X wave can be seen in the ionogram (corresponding to the plus and minus sign in Equation (2.2)). The electron density profile (black line) in the ionogram is automatically calculated following the algorithm in *Reinisch and Huang* [1983]. The profile shape can be compared with the regions of the ionosphere that have been described previously, i.e. with Figure 1. For further discussion on ionograms and how to read them see the “URSI Handbook of Ionogram Interpretation and Reduction” [Piggott and Rawer, 1978].

Lowell Digisonde International (LDI), based at Lowell, MA, are a leading constructor of ionosondes (the Digisonde). There are currently 65 operational Digisonde stations with a further 23 planned. Ionosonde data is valuable in many data assimilation models, due to the height information which can be provided. Figure 4, in comparison with Figure 2, shows that there are far fewer accessible Digisonde stations compared to GNSS stations. Although they are spread globally, Australia, central Africa and central Russia have very few stations. This can have a serious impact on global ionospheric modelling [McNamara, 2010].



Figure 4. Existing and planned ionosonde stations by Digisonde [Lowell Digisonde International, 2014b].

3. IONOSPHERIC MODELS

I have no data yet. It is a capital mistake to theorise before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts.

- Sherlock Holmes

Even though the ionosphere has been studied for many years, ionospheric modelling is an active area of research. This is because better models support the successful operation, planning and management of RF systems [Goodman and Aarons, 1990]. There are three main approaches to modelling the ionosphere: empirical, physics-based and data assimilation. Each of the approaches, and examples of each type, are described in the following sections.

A more detailed discussion of ionospheric empirical models can be found in *Bilitza* [1990], *Anderson et al.* [1989], *Bilitza et al.* [1988] and *Llewellyn and Bent* [1973]; physics-based models in *Schunk* [1996]; and data assimilation models in *Bust and Mitchell* [2008].

3.1. Empirical Models

Empirical models (also often referred to as ‘statistical models’ or ‘median models’) are based upon observations, and, as such, they provide little information about the physics of the underlying system. Relationships between the variables in the model are usually determined via curve fitting techniques [*Llewellyn and Bent*, 1973; *Bilitza et al.*, 1988; *Anderson et al.*, 1989].

Although empirical models can be used to provide information about the relationships between model variables, they cannot, on their own, be used to determine whether one variable affects

another. Also, without accounting for any of the underlying physical processes, they only allow for simple extrapolation into areas where there are no available observations.

However, compared to physics models they are much easier to construct since knowledge of the underlying physics is not required. They often provide good median performance, and are quick and computationally cheap to run. The empirical models used in this work are the International Reference Ionosphere (IRI), NeQuick and NRLMSISE-00, described in the following sections.

Other ionospheric empirical models exist which are not described in this work. Such models include the Parameterized Ionospheric Model (PIM) developed by the Air Force Research Laboratory (AFRL) [*Daniell et al.*, 1995] and the Bent Ionospheric Model [*Llewellyn and Bent*, 1973]. IRI and the Bent model have previously been compared by *Bilitza et al.* [1988]. Empirical thermospheric models include the Drag Temperature Model (DTM) [*Bruinsma et al.*, 2012] and the Jacchia Reference Atmosphere [*Jacchia*, 1977].

3.1.1. International Reference Ionosphere (IRI)

The International Reference Ionosphere (IRI) is the international standard (standard TS16457 [*ISO*, 2009]) for the terrestrial ionosphere. It is an international project sponsored by the Committee on Space Research (COSPAR) and the International Union of Radio Science (URSI) [*Bilitza*, 2004a]. IRI is an empirical model which uses the statistics of the monthly median values of the ionosphere as its basis. Such statistics include the URSI or Consultative Committee on International Radio (CCIR) world maps of foF2 and M(3000)F2 (the ratio of the maximum usable frequency at 3000 km to foF2). The model also uses plasma frequency profiles

to produce values of foF2 and hmF2. Recent improvements in the model have come from the inclusion of new ionospheric data, and better descriptions of the global and temporal variations of the ionosphere [Bilitza, 2006].

The model has a daytime height range of 65 – 2000 km and a nighttime range of 80 – 2000 km. However, the topside of the model (above foF2) is known to overestimate the electron density [Ezquer *et al.*, 1998; Bilitza and Williamson, 1999; Triskova *et al.*, 2002]. Bilitza [2004b] suggested correction factors for IRI topside based on altitude, latitude and local time. This problem can be further overcome by replacing the topside with the more accurate NeQuick topside [Coisson *et al.*, 2009], described in the next section.

3.1.2. NeQuick

NeQuick 2 is an ionospheric electron density model developed at the Aeronomy and Radiopropagation Laboratory (now Telecommunications/ICT for Development Laboratory) of the Abdus Salam International Centre for Theoretical Physics (ICTP), Trieste, Italy, and at the Institute for Geophysics, Astrophysics and Meteorology (IGAM) of the University of Graz, Austria [Nava *et al.*, 2008]. It is a median model of the ionosphere and the International Telecommunication Union Radiocommunication Sector (ITU-R) standard. It has been designed to have continuously integrable vertical profiles which allows for rapid calculation of the TEC for trans-ionospheric propagation applications. Between 100 km and the peak of the F2 layer, NeQuick uses an electron density profile based on five semi-Epstein layers [Epstein, 1930; Rawer, 1983] with modelled thickness parameters. Three profile anchor points are used: the E layer peak, the F1 peak (if present) and the F2 peak. The anchor points are defined in terms of the standard ionosonde parameters foE, foF1, foF2 and M(3000)F2 and the URSI coefficients

are often used. A semi-Epstein layer represents the model topside with a height-dependent thickness parameter that has been empirically determined. NeQuick can be run up to a height of 20000 km, and is used in the Galileo GNSS system to calculate ionospheric corrections [Angrisano *et al.*, 2013].

3.1.3. NRLMSISE-00

The US Naval Research Laboratory Mass Spectrometer and Incoherent Scatter radar Exosphere 2000 (NRLMSISE-00), is a global, empirical model of the atmosphere. It uses the 81 day average of F10.7 (the solar flux at a wavelength of 10.7 cm (2800 MHz) at the Earth's orbit – used as a proxy for solar output [Wright, 2003]), the daily F10.7 solar flux value of the previous day, and the daily Ap (which indicates the severity of the magnetic disturbances in near-Earth space [Wright, 2003]) to model the density and temperature of a number of atmospheric components [Picone *et al.*, 2002]. It is based on the earlier MSIS-86 (Mass Spectrometer and Incoherent Scatter radar 1986) [Hedin, 1987] and MSISE-90 (Mass Spectrometer and Incoherent Scatter radar Exosphere 1990) [Hedin, 1991] models.

The model outputs number densities of helium, atomic oxygen, molecular oxygen, atomic nitrogen, molecular nitrogen, hydrogen and argon, as well as total mass density and the temperature at a given altitude. NRLMSISE-00 has been shown to offer a noticeable improvement over MSISE-90 [Picone *et al.*, 2002] and Jacchia-70 [Jacchia, 1977]. Although it is not an ionospheric model, and does not output electron density, it is included here as it is required by the physics-based models described in Section 3.2.

3.2. Physics-Based Models

Physics-based models solve the equations which describe the physical processes in the ionosphere/thermosphere. To model atmospheric density the models generally use empirical models (such as NRLMSISE-00) to generate an initial condition. Neutral and ion species production is then calculated via chemical reaction equations and using solar X-rays and EUV conditions. Ion transportation and recombination are also considered. The initial and boundary conditions, as well as proxies for solar activity, are the main drivers for the models. There are a number of approaches to modelling the physics of the ionosphere, which rely on different numerical methods [Purnell, 1976; Augenbaum, 1984; Bott, 1989], and thus exhibit different levels of complexity and use a variety of inputs.

Physics models have the potential to provide ionospheric forecasts since, in principle, the physics can be used to propagate the model densities from one time step to the next. However, Shim *et al.* [2012] suggested that errors in electron density can be very large due to errors in initialization and boundary conditions. This renders them ineffective for real-time services. Generally, whilst they accurately show the effect of varying input conditions, it is not easy to provide accurate absolute values of the ionospheric parameters unless one has an accurate starting point for the modelling [Shim *et al.*, 2012].

The physics-based models used in this work are TIE-GCM and GITM, which are described in the following sections. Other thermosphere/ionosphere physics-based models exist which are not described here. Such models include the Coupled Thermosphere Ionosphere Plasmasphere Electrodynamics Model (CTIPE) developed at the Space Weather Prediction Center [Codrescu

et al., 2012] and the Coupled Middle Atmosphere and Thermosphere (CMAT) model from University College London [*Harris*, 2011].

3.2.1. TIE-GCM

The National Center for Atmospheric Research (NCAR) Thermosphere Ionosphere Exosphere – General Circulation Model (TIE-GCM) is a three-dimensional model of the coupled thermosphere ionosphere system [*Richmond*, 1992]. At each time step the continuity, energy and momentum equations are solved for neutral and ion species using a fourth-order, centred finite difference scheme [*Roble et al.*, 1988]. TIE-GCM has a user selected latitude and longitude grid, and 29 constant-pressure levels, which range in height from approximately 97 km to 500 km.

The model takes as input the daily F10.7, the 81 day F10.7 average and the Ap. It uses either the Weimer or Heelis models for the ionospheric electric fields at high latitudes [*Heelis et al.*, 1982; *Weimer*, 2005]. Throughout this work, the Heelis model has been used. The lower boundary condition (atmospheric tides) are given by the Global Scale Wave Model (GSWM) [*Hagan et al.*, 1999]. *Qian et al.* [2009] has shown that TIE-GCMs model errors are dependent on the quality of its input parameters.

3.2.2. GITM

The Global Ionosphere Thermosphere Model (GITM) is a physics-based three-dimensional global model that solves the full Navier-Stokes equations for density, velocity, and temperature for a number of neutral and charged components [*Ridley et al.*, 2006]. The model also provides

the total neutral density, electron density, electron, ion and neutral temperatures, neutral wind speed and plasma velocities. For inputs, GITM uses F10.7 solar flux, hemispheric power index (HPI) [Emery *et al.*, 2008] (derived from the 3-hour Kp (a proxy for geomagnetic activity, similar to Ap [Wright, 2003])), interplanetary magnetic field (IMF) data and solar wind velocity. The HPI file format was changed in 2007 [U.S. Dept. of Commerce, NOAA, 2007] and GITM requires the pre-2007 format. Data files after this date must be converted to the old format before they can be used as input. The model allows the user to select latitude and longitude grids and uses a fixed altitude grid for the height profile.

To solve the continuity, energy and momentum equations, GITM uses an advection solver, whilst the ion momentum equation is solved assuming a steady state [Ridley *et al.*, 2006]. The initial state can be set in three ways: either using an ideal atmosphere (the densities and temperature at the bottom of the atmosphere are passed into the model); using NRLMSISE-00 and IRI, or by using a previous run. GITM inherently allows for non-hydrostatic solutions to develop which allows for realistic dynamics in the auroral zones [Ridley *et al.*, 2006].

3.3. Data Assimilation

As discussed in Section 2.7, there is a wide range of ionospheric data available. This data enables a further modelling approach – data assimilation. Data assimilation models aim to optimally combine disparate measurements with a background model. A range of assimilation methods have been developed; i.e. weighted least squares (WLS) [Plackett, 1950], Kalman filters [Kalman, 1960; Houtekamer and Mitchell, 2005], Optimal Interpolation (OI) [Gandin, 1963; Eddy, 1967] and variational methods [Le Dimet and Talagrand, 1986]. The background model may, in principle, be empirical or physics-based. Data assimilation models are not real-

time, since there is some delay between the data being retrieved and it being assimilated, but they can work in ‘near-real-time’ [Angling and Khattatov, 2006]. A review of ionospheric data assimilation models can be found in *Bust and Mitchell* [2008].

In many systems there are usually fewer observations than the number of unknowns, i.e. the system is mathematically underdetermined. The observations which are available are subject to experimental uncertainty, and often the observed characteristic is not the variable that is directly modelled. Therefore an observation operator (\mathbf{H}) is used to transform the state space to the observation space, i.e.

$$\mathbf{y}_0 = \mathbf{H}\mathbf{x} + \boldsymbol{\epsilon}_0. \quad (3.1)$$

Where \mathbf{y}_0 is a vector of observations, \mathbf{x} the true state and $\boldsymbol{\epsilon}_0$ the observation error.

It is common to want to estimate the state of a system from some given information; for example, to find the TEC through a particular part of the ionosphere. Ideally, a satellite and receiver would be placed at suitable locations, and the TEC directly measured, albeit with some measurement error. However, in practice, this is not possible and the value has to be estimated.

The first approach is to just use data from the nearest satellite/receiver pair to the region of interest. Secondly, if more than one satellite/receiver pair exist, then both can be used to interpolate the information in order to estimate the TEC. This estimation can be improved by using more satellite/receiver data. However, depending on a number of factors, including the distance of the satellite from the point of interest, each new piece of information should be

weighted differently depending on how much the data is “trusted”. Data assimilation is the art of optimally combining various sources of information to estimate the true state of the system.

The subject of data assimilation is both wide and varied. The short introduction to the topic which follows merely serves to highlight the essence of the field, with a primary focus on Kalman filters. There are a number of excellent books on the subject for further reading, specifically *Rodgers* [2000], *Kalnay* [2003] and *Evensen* [2009].

3.3.1. Least Squares

A simple approach to data assimilation is to use the method of least squares [*Legendre*, 1805].

The method requires the estimation of the state to be a value that minimises the sum of the squares of the error of each term. That is, for n observations, Ω_i for $i = 1, n$, a variable x_a such that:

$$\sum_{i=1}^N (x_a - \Omega_i)^2 \text{ is a minimum.} \quad (3.2)$$

This approach is often used for data fitting (constructing curves of best fit to a series of data).

It is not used in data assimilation models since, in its classic form, does not allow for differently weighted data. This means that the result can easily be skewed by outliers.

3.3.2. Weighted Least Squares

To overcome the problem with least squares, discussed in the previous section, a somewhat more sophisticated data assimilation technique can be used, weighted least squares (WLS). Assuming the data is unbiased, the WLS is formed from a linear combination of observations and has the minimum variance of all such combinations. To derive the WLS consider a linear combination of two vectors of data, \mathbf{x}_b (the “background”, so named for consistency in other works) and \mathbf{y}_0 (the “observation”), to produce the “analysis”, \mathbf{x}_a :

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y}_0 - \mathbf{x}_b), \quad (3.3)$$

where \mathbf{K} is the so called Kalman gain matrix, and \mathbf{x}_a , \mathbf{x}_b and \mathbf{y}_0 have associated error covariance matrices \mathbf{A} , \mathbf{B} and \mathbf{O} . Equations (3.3) and (3.4) are examples of the equations without the observation operator. That is when the number of observations and size of background matrix is the same, and the state and observation spaces are the same. To find the WLS, a matrix \mathbf{K} must be found which minimises the expected error covariance matrix \mathbf{A} associated with \mathbf{x}_a . This is equivalent to minimising the trace of \mathbf{K} [Park and Xu, 2013], and is given by:

$$\mathbf{K} = \mathbf{B}(\mathbf{B} + \mathbf{O})^{-1}. \quad (3.4)$$

A detailed description of the derivation of Equation (3.4) can be found in Kelly [2013].

Using the observation operator, as described in Section 3.3, the WLS for vector equations can be transformed from the observation space to the state space. For example Equation (3.4) becomes:

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{O})^{-1}. \quad (3.5)$$

It is in this form that the equations are usually presented.

If the conditions of the WLS are restricted further, i.e. that the observation errors (residuals) are unbiased and have the same variance, then it can be shown to be the best linear unbiased estimate (BLUE) [Plackett, 1950].

3.3.3. Kalman Filters

Finding the WLS for a variable can be further developed by replacing the unique state \mathbf{x} with a series of states \mathbf{x}^t , where t is a time index. There are then a number of methods which one can use to solve these equations and create the best analysis. This includes variational assimilation, which avoids the inversion required to compute \mathbf{K} by considering the analysis as an approximate solution to the equivalent minimization problem [Le Dimet and Talagrand, 1986]. An alternative to variational assimilation is the Kalman filter which is an extension of the least-squares analysis [Kalman, 1960].

The Kalman filter is a technique used to perform data assimilation, where each background state is provided by an “update” step, which starts from the previous analysis [Kalman, 1960]. There are a number of conditions and assumptions which are required before using a Kalman filter. First, initial background data is required to start the procedure (\mathbf{x}_b^0). A (linear) model which describes the evolution of the state, i.e. an approximation of how the state which is being

modelled changes over time, is also required (\mathbf{M}^t). It need not be a perfect description of the evolution of the state, since it is combined with data. Finally there are some assumptions made about the errors in the system. The errors (model and observation) are assumed to have no bias, i.e. $E[\boldsymbol{\eta}^t] = 0$ and $E[\boldsymbol{\epsilon}_i^t] = 0$, where $\boldsymbol{\eta}^t$ and $\boldsymbol{\epsilon}_i^t$ are the model and observation errors respectively. Also, different types of errors, such as the model and observation errors, are independent of each other, $E[\boldsymbol{\eta}^t \boldsymbol{\epsilon}_i^t] = 0$.

Given a series of states, \mathbf{x}^t , the Kalman filter will give the optimal (i.e. the minimum mean square error) estimate of the state, \mathbf{x}_a^t , using observations $\mathbf{y}_o^1, \mathbf{y}_o^2, \dots, \mathbf{y}_o^t$. The filter works in two, repeating, steps:

1. Analysis,
2. Update.

The initial conditions, \mathbf{x}_b^0 and \mathbf{B}^0 , are the starting state and its corresponding error covariance matrix. These are then used to perform the analysis step:

Analysis step. This uses the vector equations for the WLS with the observation operator (Equations (3.3) and (3.5)) as well as the error covariance matrix [Kelly, 2013] giving:

$$\mathbf{K}^t = \mathbf{B}^t \mathbf{H}^T (\mathbf{H} \mathbf{B}^t \mathbf{H}^T + \mathbf{O})^{-1}, \quad (3.6)$$

$$\mathbf{x}_a^t = \mathbf{x}_b^t + \mathbf{K} (\mathbf{y}_o^t - \mathbf{H} \mathbf{x}_b^t), \quad (3.7)$$

$$\mathbf{A}^t = (\mathbf{I} - \mathbf{K}^t \mathbf{H}) \mathbf{B}^t. \quad (3.8)$$

Where \mathbf{I} is the identity matrix.

Update step. Using the values from the analysis step, the update step is used to find the background values \mathbf{x}_b^{t+1} and \mathbf{B}^{t+1} for use in the next analysis step.

$$\mathbf{x}_b^{t+1} = \mathbf{M}^{t,t+1} \mathbf{x}_a^t, \quad (3.9)$$

$$\mathbf{B}^{t+1} = \mathbf{M}^{t,t+1} \mathbf{A} (\mathbf{M}^{t,t+1})^T + \mathbf{Q}^t. \quad (3.10)$$

Where \mathbf{Q}^t is the model error covariance matrix. This process is repeated t times, until the assimilation is complete.

When using a Kalman filter the error covariance matrices scale with the square of the size of the state. For large states this makes them difficult to work with, especially in a reasonable time scale. This is the main problem with the Kalman filter. Although it provides the best combination of the modelled data it is computationally expensive. To overcome this, and other problems (such as the propagation model needing to be linear), a number of variants of the Kalman filter have been proposed. Such variants include: the band limited Kalman filter, which only saves part of the covariance matrix based on physical correlation lengths [*Hajj et al.*, 2004]; the extended Kalman filter, which allows the use of non-linear models in the Kalman filter [*Julier and Uhlmann*, 2004]; the ensemble Kalman filter, which replaces the covariance matrix by the sample covariance [*Evensen*, 1994] and the local ensemble Kalman filter which performs the analysis around each model grid point [*Ott et al.*, 2004].

It is apparent that there is a wide range of techniques for data assimilation. This is further reflected by the range of models which are based upon assimilative routines. The assimilation

models used in this work are EDAM, GPSII and the data ingestion version of NeQuick. Other data assimilation models, which are not discussed here, include the Utah State University Global Assimilation of Ionospheric Measurements (USU-GAIM) Model [McNamara *et al.*, 2008] and the Jet Propulsion Laboratory's Global Assimilative Ionospheric Model (GAIM) [Mandrake *et al.*, 2005] which both use data assimilation techniques on top of a physics-based background model. Data assimilation is also used in a 3D tomography scheme by the Multi-Instrument Data Analysis System (MIDAS) [Mitchell and Spencer, 2003; Spencer and Mitchell, 2007].

3.3.4. Electron Density Assimilative Model (EDAM)

The Electron Density Assimilative Model (EDAM) has been developed at QinetiQ, UK to assimilate ionospheric measurements into a background ionospheric model [Angling and Khattatov, 2006; Angling *et al.*, 2009]. The background model is provided by IRI-2007 (the 2007 version of IRI, Section 3.1.1) [Bilitza and Reinisch, 2008]. The EDAM assimilation is based on the weighted least squares described in Section 3.3.2.

A magnetic coordinate system (tilted dipole) is used in EDAM and a time step of 15 minutes. The electron density differences between the voxels of the analysis and the background model are propagated from one time step to the next by assuming persistence combined with an exponential decay. The time constant for this decay is set at four hours. Thus if the data feed is interrupted, the analysis will decay back to the background model.

EDAM assimilates GNSS slant TEC observations as well as ionosonde information in the form of the ionogram virtual height traces [Angling and Jackson-Booth, 2011]. It can also assimilate

radio occultation (RO) observations of slant TEC [*Angling and Cannon, 2004; Angling, 2008*] and in situ electron densities, but these capabilities were not used in this work.

3.3.5. GPS Ionospheric Inversion (GPSII)

The GPS Ionospheric Inversion (GPSII) model uses a recursive data assimilation algorithm [*Fridman et al., 2006, 2009*] that provides an electron density distribution model for a geographical area. At each time step of the solution, the resulting electron density model matches all ionospheric data accumulated during the model update time interval to within the data measurement error. This data-driven modification of the model is determined with the Tikhonov method [*Tikhonov and Arsenin, 1977*]. The Tikhonov method ensures that the resulting model is spatially and temporally smooth.

GPSII is able to assimilate absolute and relative TEC data from ground- and space-based GNSS receivers, relative TEC data obtained with low Earth orbit (LEO) satellite beacons, in-situ electron density measurements and data from vertical incidence ionosondes. GPSII is also able to assimilate various HF measurements such as OTHR backscatter ionograms [*Fridman et al., 2012*], but these capabilities were not used in this work. GPSII also uses IRI-2007 [*Bilitza and Reinisch, 2008*] (Section 3.1.1) as the background model and the solution update interval is 15 minutes.

3.3.6. NeQuick Data Ingestion Model

The empirical NeQuick model was described in Section 3.1.2. However, to provide real time 3D specifications of the ionospheric electron density, the NeQuick output can be modified using

GNSS derived TEC data and profile peak parameters measured by ionosonde values [Nava *et al.*, 2006; Buresova *et al.*, 2009]. NeQuick uses an effective ionisation level (A_z) as a measure of the level of solar activity. In the real time system, A_z_{foF2} is defined to be the effective ionization level that allows NeQuick to reproduce the observed foF2 value at a reference station. Then via the Dudeney formula [Dudeney, 1978], an A_z_{hmF2} value is specified that allows NeQuick to reproduce the observed hmF2. These parameters, A_z_{foF2} and A_z_{hmF2} , are used to locally constrain the model peak parameters. The TEC error, as compared to the observation, is then minimised by modifying the model's bottomside thickness (i.e. the slab thickness is varied). The three effective parameters (A_z_{foF2} , A_z_{hmF2} and the slab thickness correction factor), are then used to run the model over the area of interest.

This method relies on available and valid data from an ionosonde. Consequently, the data ingestion version of the NeQuick model does not provide any output when there is no such ionosonde data. This version of the model will be referred to herein as mNeQuick.

3.4. Multi-Model Ensemble Introduction

The uncertainties in any given model include incomplete physics, errors in initial conditions, boundary conditions and parameter values. By combining models together, to form a multi-model ensemble (MME), it is possible to reduce the impact of these errors and hence increase the model forecast skill [Tebaldi and Knutti, 2007]. MMEs are widely used in weather and climate modelling, where there have been multiple approaches proposed for constructing the ensemble [Krishnamurti *et al.*, 2000; Kharin and Zwiers, 2002].

Tebaldi and Knutti [2007] argue that the expectation that combining results from multiple models will offer an improvement is based on the assumption that, if the models are independent, errors should cancel. This leads to the idea that the uncertainty in the model predictions should decrease as the number of models increases. However, it has been shown that different weighting schemes, each based on reasonable criteria, can give different results [*Hagedorn et al.*, 2005]. Therefore, *Chandler* [2013] concluded that, without a formal reasoning for a particular weighting method, the choice becomes an added source of uncertainty. Thus, simply increasing the number of models in the MME should not continually decrease the uncertainty without formally characterising a method for combining the models [*Tebaldi and Knutti*, 2007].

3.5. Conclusions

There are a variety of modelling techniques used to describe the ionosphere including: empirical, physics-based and data assimilation. These techniques have been described, as have associated models. MMEs have also been introduced, which are further described and used in Chapter 6.

Before models can be used operationally in RF systems their performance should be measured. Relative model performance however is difficult to measure due to the metrics used to quantify performance. There are a variety of metrics which can be used to measure the model skill. Also, different model parameters will give different skill scores. How this combination of factors to judge model performance should be balanced is an important aspect of model comparison. The following chapter describes one such approach to model performance, modified Taylor diagrams.

4. HOW TO COMPARE IONOSPHERIC MODELS

*He uses statistics as a drunken man uses lamp posts;
for support rather than illumination.*

- Andrew Lang

This chapter is an extended version of a paper published in Radio Science [Elvidge *et al.*, 2014].

4.1. Current Model Comparison Techniques

Models are used throughout the scientific disciplines, and the comparison of them is critical in their evaluation. Intuitively, it seems that the best way to compare two models is to see how closely the models reproduce an observation. However, models usually have multiple parameters, which themselves are not always directly measurable, and therefore the task is often not straightforward. Nonetheless, a number of sophisticated techniques for model comparisons have been developed and are used in a variety of fields. Such methods include Bayesian model selection [Kass and Raftery, 1996], minimum description length [Rissanen, 1983], Akaike and Bayesian Information Criterion [Akaike, 1973; Schwarz, 1978] and cross-validation [Stone, 1974].

Ionospheric model comparison, though, is still mostly accomplished via more basic techniques. In general, ionospheric models are usually compared, parameter by parameter, directly with observations. Means, standard deviation of differences and the root mean square (RMS) errors are commonly compared for individual model parameters.

McNamara *et al.* [2013] compared models by considering the RMS errors, means and standard deviation values, separately, for a month long test period. In one case the analysis was split between day and night time results. Other examples include Shim *et al.* [2011] who compare model performance by using four different metrics: the RMS error, prediction efficiency (comparing the ratio of the modelled RMS error and of a reference model RMS difference (in this case the mean value of the observation)), ratio of the maximum change in amplitude and the ratio of the maximum amplitude. Angling and Khattatov [2006] compared two ionospheric models by again investigating the RMS error, the mean error and by discussing the time series plots of the tested parameters (foF2 and hmF2).

How to present such results is also a challenge. Time series figures of a particular parameter from multiple models can be confusing, and fail to easily deliver the required information (Figure 5). The alternative approach is to either separate similar figures into individual plots, or list the results in tables; neither provide satisfactory solutions.

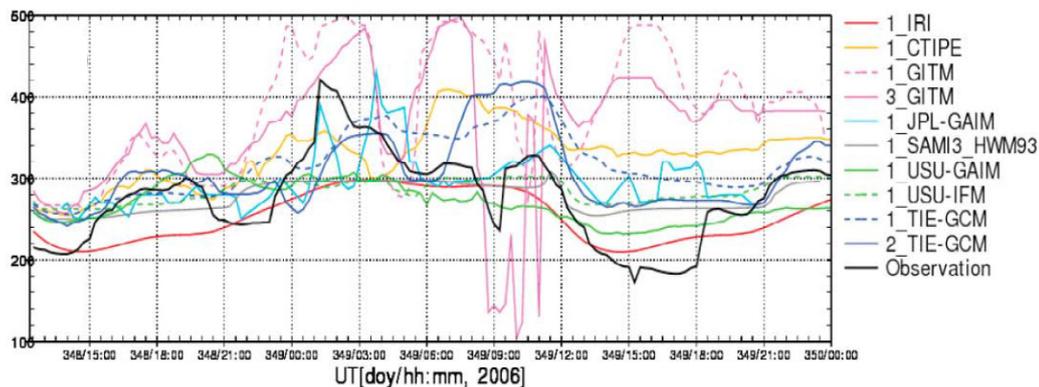


Figure 5. Observed (black curve) and modelled (colour curves) of hmF2 at a particular testing site. Reproduced from Shim *et al.* [2011].

4.2. Taylor Diagrams

Taylor diagrams can be used to graphically summarise how closely a set of patterns (in this case a set of models) match observations [Taylor, 2001]. This is achieved by plotting the standard deviation of the time series of each model parameter and the correlation (the Pearson product-moment correlation) between the time series of the model values and the observations. Taylor also noted the geometric connection between correlation, standard deviation and the centred-pattern RMS difference. The centred-pattern RMS difference is the mean removed RMS difference, calculated thus:

$$E^2 = \frac{1}{N} \sum_{n=1}^N [(m_n - \bar{m}) - (d_n - \bar{d})]^2, \quad (4.1)$$

where E is the centred-pattern RMS difference, m_n is the time series of the model parameter that is being tested and d_n is the time series of the observation data. These statistics are related by

$$E^2 = \sigma_m^2 + \sigma_d^2 - 2\sigma_m\sigma_d R. \quad (4.2)$$

Where σ_m and σ_d are the standard deviations of the model and observation respectively and R is the correlation coefficient. E , R , σ_m and σ_d can then be plotted on the same diagram.

Care should be taken to distinguish between the RMS difference and centred-pattern RMS difference as used in Taylor diagrams. The centred-pattern RMS difference is a mean-removed RMS difference and cannot be equated with the RMS difference. Although in the literature,

where Taylor diagrams are used, it often is; for example *Freidenreich et al.* [2004], *AchutaRao and Sperber* [2006], *Miller et al.* [2006], *Maraun et al.* [2010] and *Mote and Salathe* [2010]. If two models have no bias, then a reduction in centred-pattern RMS difference implies a reduction in the RMS difference. So, with this assumption, works which erroneously state there is a reduction in RMS difference, when in fact the Taylor diagram shows there is only a reduction in centred-pattern RMS difference, are valid. However this assumption fails if either model has any bias.

Taylor diagrams address one of the problems discussed in Section 4.1; that of comparing only one parameter at a time. The diagrams allow different parameters to be plotted together by normalising (i.e. making dimensionless) their standard deviations. This allows for visualization of multiple parameters simultaneously, giving a better overall picture of the models performance. The diagrams are often used in the study of climatology, and their associated models [*Maurer et al.*, 2002; *Chang and Hanna*, 2004; *Gleckler et al.*, 2008; *Kobayashi et al.*, 2014].

Figure 6 is an example of a Taylor diagram.

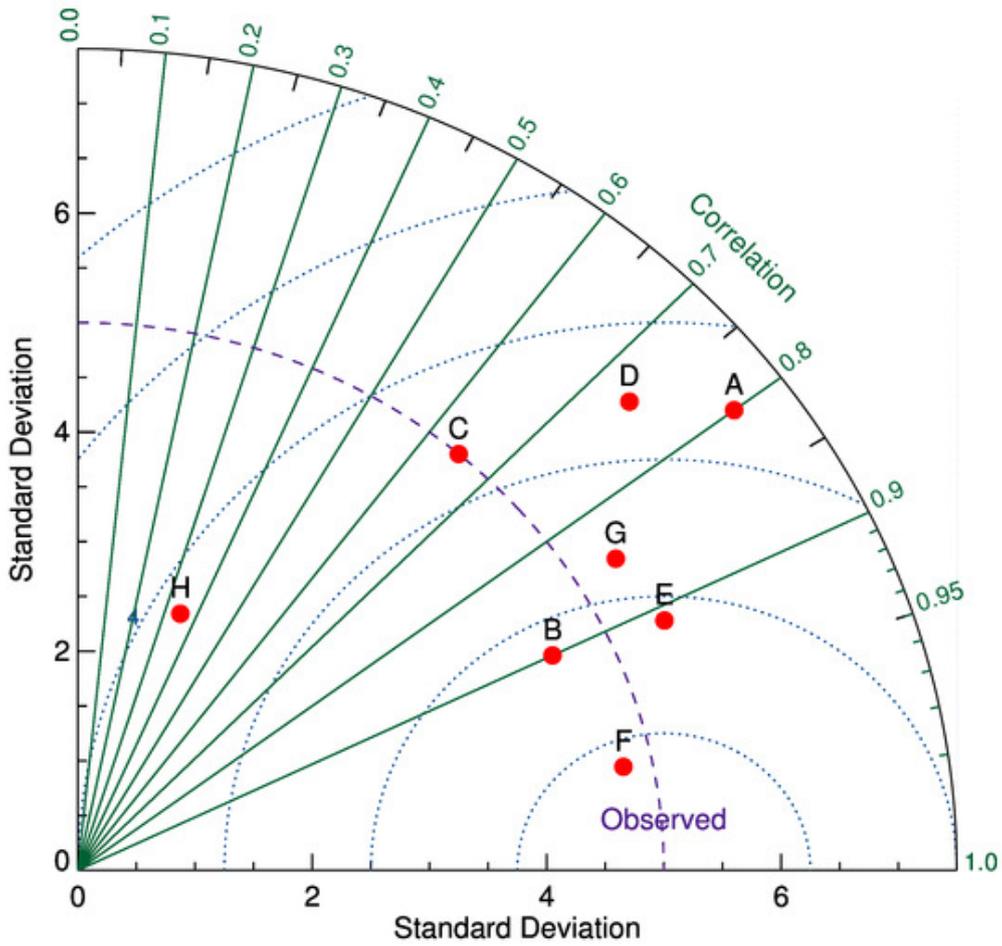


Figure 6. Example Taylor diagram for models 'A' through 'H'. The dashed line (labelled observed), shows the standard deviation of the observations. The dotted line semi-circles, originating from the intersection of the observed standard deviation (dashed line) and the horizontal axis, show the centred pattern RMS value. Considering a specific model, the radial distance from the origin shows the models' standard deviations and the azimuthal angle corresponds to the correlation between the model and the observation.

4.3. Modified Taylor Diagrams

Many of the characteristics of Taylor diagrams are extremely useful for ionospheric model comparison. However they do have the confusion of centred-pattern RMS difference and no way of showing model bias. Presented here are modifications to such diagrams to address these issues.

It can be shown (Equations (4.3) to (4.5)), that the centred-pattern RMS difference is equal to the standard deviation of the model error (model minus observation).

$$E^2 = \frac{1}{N} \sum_{n=1}^N [(m_n - \bar{m}) - (d_n - \bar{d})]^2, \quad (4.3)$$

$$= \frac{1}{N} \sum_{n=1}^N [(m_n - d_n) - (\bar{m} - \bar{d})]^2, \quad (4.4)$$

$$= \frac{1}{N} \sum_{n=1}^N [(m_n - d_n) - (\overline{m - d})]^2. \quad (4.5)$$

The standard deviation of the model error is a more familiar term in model testing, and is used in place of centred-pattern RMS difference henceforth.

Bias is an important feature in modelling, and the ability to show this on the diagram would be useful. One approach suggested by *Taylor* [2001] is to add lines, whose length is equal to the bias, to each data point. An alternative approach, used as a modification here, is to show the bias of the models via a colour scale. Naturally, the bias, like the standard deviation, must also be normalized in order to plot multiple parameters on a single diagram. It should also be noted

that the mean square difference between a model and the data can be calculated by adding in quadrature the bias and the standard deviation of the errors.

Figure 7 shows an example Taylor diagram, with the modifications described in this section (herein referred to as ‘modified Taylor diagrams’). The radial distance of a data point from the origin is a model’s normalised standard deviation (in this instance normalised by the standard deviation of the observation) and the azimuthal angle corresponds to the correlation between the model and observation. The dashed line (labelled ‘Observed’), shows the normalised standard deviation of the observation (i.e. unity). The dotted-lined semi-circles, originating from the intersection of the observed standard deviation (dashed line) and the horizontal axis, show contours of the standard deviation of the model error. The model bias is shown with a colour scale and the mean square error of the model can be found by adding in quadrature the bias and standard deviation of the errors. Different symbols represent different models, whereas the labels represent different parameters. In this way models can be compared for a range of parameters simultaneously. Finally, the normalisation factor, required to retrieve the original parameter values, is shown in the top right.

Figure 7 is a modified Taylor diagram for three different models, with four parameters, presenting five different pieces of statistical information. A table to represent the same quantity of data would require 60 entries.

4.4. Conclusions

Taylor diagrams have been modified in order to show model bias, standard deviation of errors and a method of calculating the mean square error. In this chapter modified Taylor diagrams

have been presented and described as a method for ionospheric model comparison. They provide an easy and intuitive way of presenting multiple statistics for a range of parameters when comparing a model to observation data.

Modified Taylor diagrams are used in Chapter 5 for a series of ionospheric model comparisons.

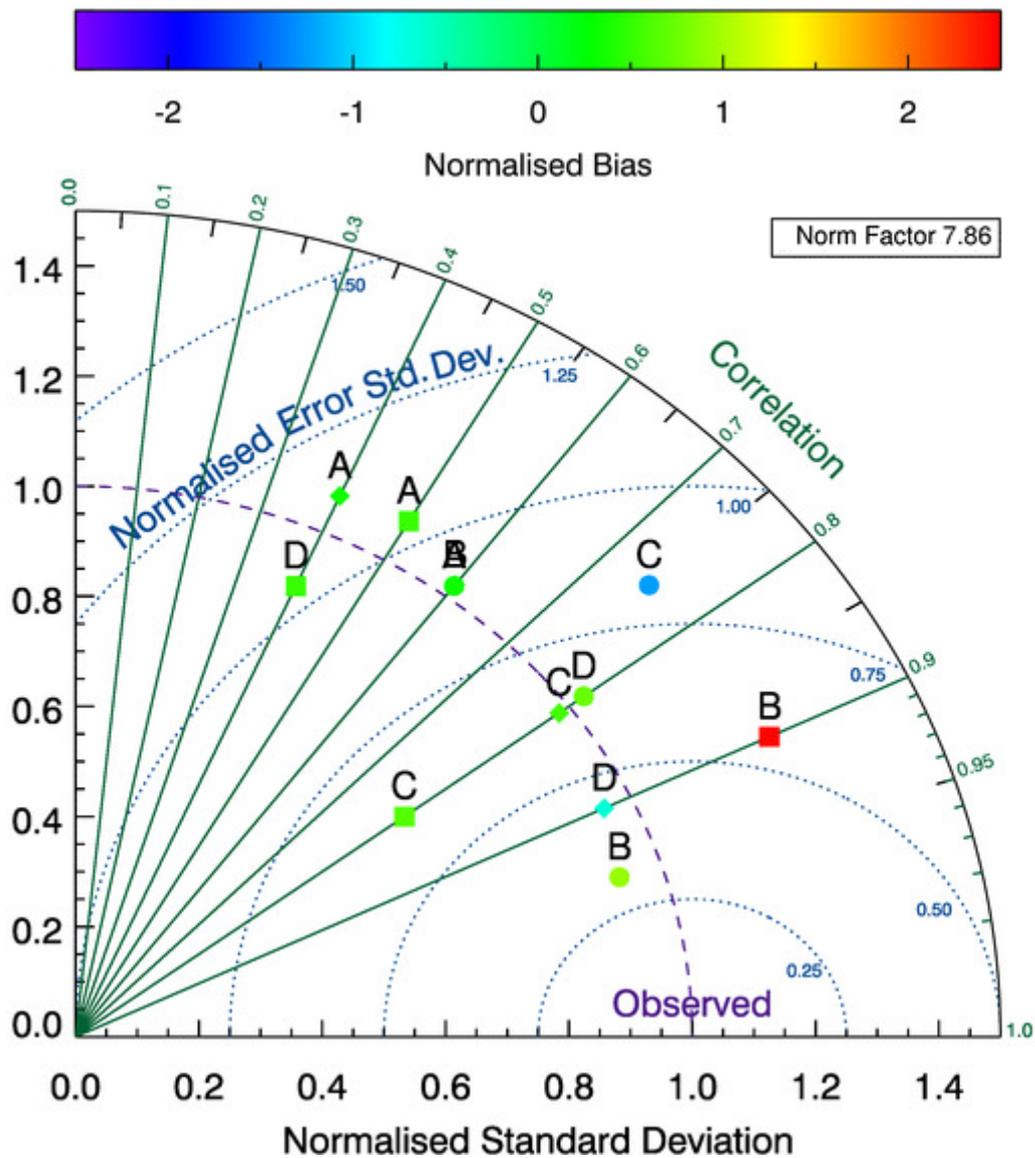


Figure 7. Example modified Taylor diagram.

5. MODEL COMPARISONS

To compare is not to prove.

- French Proverb

The work presented here is an extended version of that published in *McNamara et al.* [2013] and *Elvidge and Angling* [2014].

As well as metric(s) to measure model performance, such as those presented by modified Taylor diagrams described in Section 4.3, common testing scenarios are required. *Shim et al.* [2011] described nine different testing scenarios and used them to compare the three main model types: empirical, physics-based and data assimilation. They compared the ability of the models to specify NmF2, hmF2 and vertical drift, with respect to data from incoherent scatter radars.

Although nine scenarios were used, they were all relatively short; ranging from one to two and a half days. The time-step was 15 minutes, and therefore a fairly small number of data points were used in the calculation of the metric statistics. This leads to two main statistical problems. First, a small number of data points can result in large confidence intervals for the estimate of the statistics. Secondly there is an increased chance that the time series may be non-stationary.

An $\alpha\%$ confidence interval, for a given parameter, is an interval in which, for a large number of samples, the probability that the confidence interval contains the parameter is $\alpha\%$. The upper and lower confidence intervals of the mean are given by:

$$\text{Confidence Interval} = \pm \frac{t^* \sigma}{\sqrt{n}}, \quad (5.1)$$

where t^* is the critical value (a t-table value depending on the confidence level required), σ is the standard deviation of the sample, n is the sample size and the positive part is for the upper confidence interval and the negative for the lower [Sheskin, 2004]. It can therefore be seen that the confidence intervals decrease as a factor of the square root of the sample size.

For estimates of the standard deviation, confidence intervals are defined as [Sheskin, 2004]:

$$\text{Lower Limit} = \sigma \times \sqrt{\frac{n-1}{\text{Inv}\chi^2\left(\frac{\alpha}{2}, n-1\right)}}, \quad (5.2)$$

$$\text{Upper Limit} = \sigma \times \sqrt{\frac{n-1}{\text{Inv}\chi^2\left(1-\frac{\alpha}{2}, n-1\right)}}. \quad (5.3)$$

Where σ is the standard deviation, n is the sample size, $\text{Inv}\chi^2$ is the inverse χ^2 distribution and α is the confidence level. Thus, the greater the sample size the closer to the standard deviation the confidence intervals are.

A comparison of empirical, physics-based and data assimilation models are presented here for a much longer test scenario than used by Shim *et al.* [2011]. The study is also part of the first extensive analysis of EDAM results where ionogram traces have been assimilated and independent ionosondes used to provide truth data. An earlier study addressed only a single day

of data and was limited to an examination of the residual errors [Angling and Jackson-Booth, 2011].

5.1. Testing Scenario

The models under test can all be used for global specification. However, the *models' regions of effective coverage are restricted to areas where data is available. This test concentrates on observations from the four ionosondes (Digisonde DPS-4D [Reinisch et al., 2009]) in the Republic of South Africa (RSA) (Figure 8 and Table 1).*

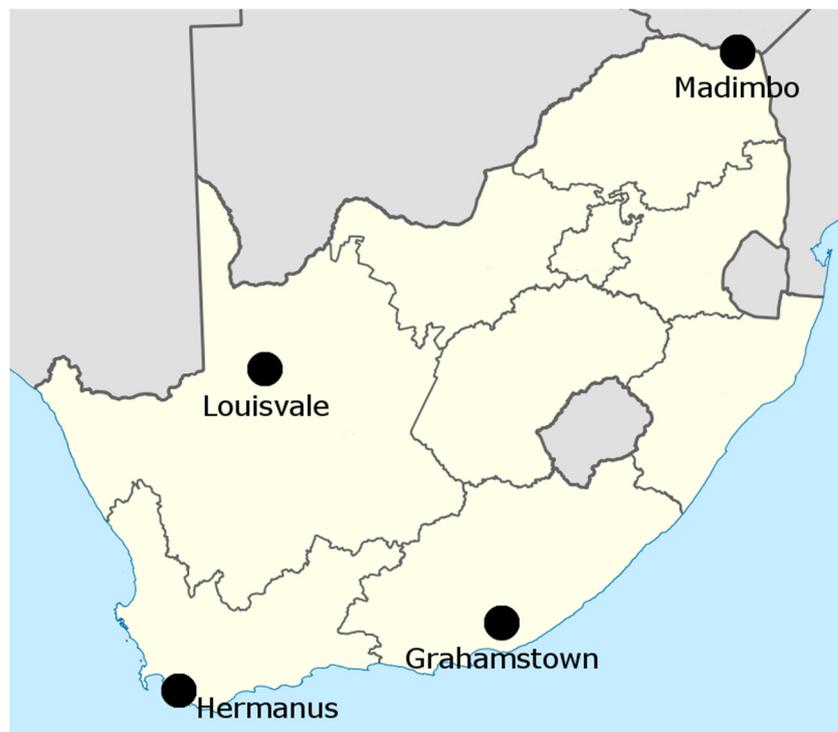


Figure 8. Locations of the four Republic of South Africa ionosondes. Louisvale and Hermanus are ~700 km from Grahamstown. Madimbo is ~1300 km from Grahamstown.

Table 1. Locations of ionosonde data assimilated in the test.

Short Name	Long Name	Latitude	Longitude
GR13L	Grahamstown	-33.3	26.5
HE13N	Hermanus	-34.4	19.2
LV12P	Louisvale	-28.5	21.2
MU12K	Madimbo	-22.4	30.9

For this study only the Grahamstown (GR13L) ionograms and sub-Saharan GNSS data were assimilated. In particular, the GNSS sites listed in Table 2 were used. Their positions are shown in Figure 9 and Table 2.

Table 2. GNSS stations assimilated in the test.

Short Name	Long Name	Latitude	Longitude
ABPO	Antananarivo	-19.0	47.2
ADIS	Addis Ababa	9.0	38.8
HARB	Pretoria	-25.9	27.7
HNUS	Hermanus	-34.4	19.2
HRAO	Hartebeesthoek	-25.9	27.7
MAL2	Malindi	-3.0	40.2
MAUA	Maun	-19.9	23.5
MFKG	Mafikeng	-25.8	25.5
NKLG	N’Koltang	0.4	9.7
NURK	Kigali	-1.9	30.1
RBAY	Richardsbay	-28.8	32.1
RCMN	Nairobi	-1.2	36.9
REUN	La Reunion	-21.2	55.6
SEY1	La Misere	-4.7	55.5
SUTH	Sutherland	-32.4	20.8
TDOU	Thohoyandou	-23.1	30.4
WIND	Windhoek	-22.6	17.1
ZAMB	Lusaka	-15.4	28.3

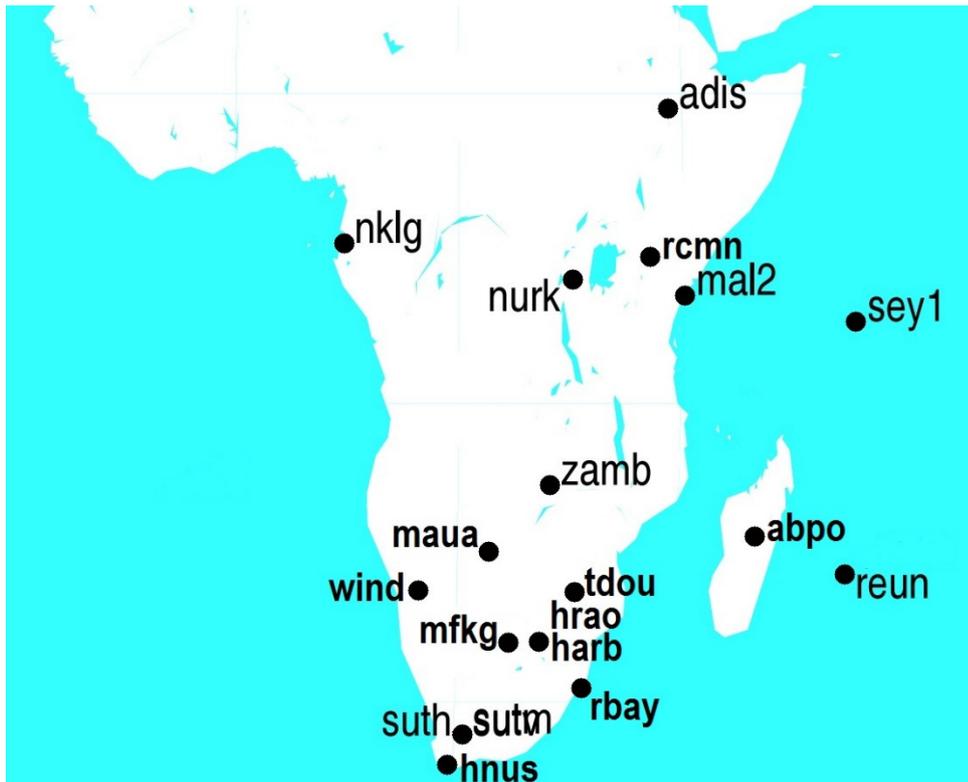


Figure 9. Location of the Sub-Saharan Africa GNSS stations used in this test.

Hermanus (HE13N), Louisvale (LV12P) and Madimbo (MU12K) ionograms were used to provide the ground truth observations. It is these observations against which the models are tested.

The whole month of September 2011 was used for the test period and each model ran with a 15 minute time step. Thus a maximum of 2880 data points are available to estimate the monthly statistics. September 2011 showed a large amount of solar and geomagnetic variability. The daily sunspot number ranged from 47 to 173 and the Ap exceeded the storm threshold (29 [Space Weather Prediction Center, 2009]) on five days (Figure 10).

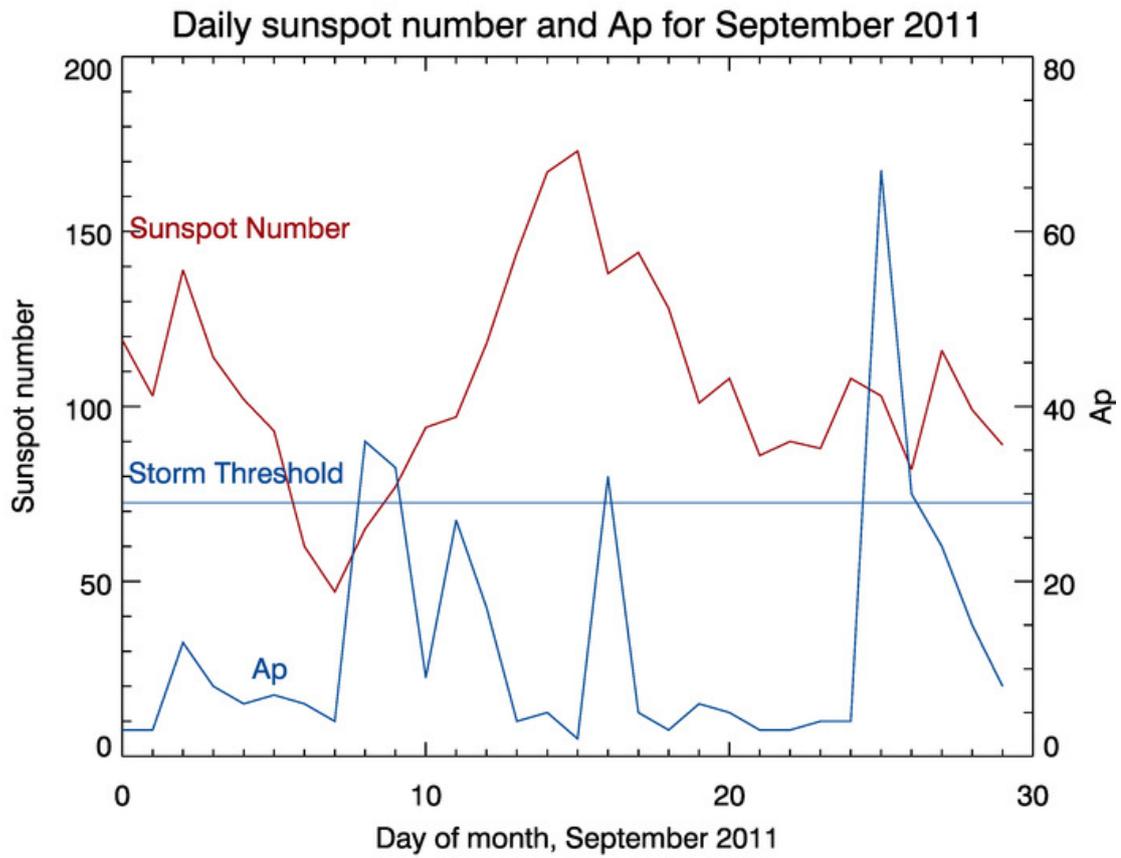


Figure 10. Time series plot of the sunspot number (red) and Ap (blue) for September 2011. Ap values above 29 exceed the storm threshold as defined by the Space Weather Prediction Center [2009].

5.2. Filtering Observations and Model Results

All the ionosonde observations were automatically processed (autoscaled) by ARTIST-5 [Galkin and Reinisch, 2008]. Although the RSA ionosondes are well maintained and ARTIST-5 is a significant advance on earlier versions of ARTIST, there are still some autoscaling mistakes. These mistakes can lead to unphysical results and procedures have been adopted to prevent them from distorting the results of the present study:

- foF2 values greater than 15 MHz produced by the ionosonde were ignored. This value was chosen as it equates to a two sigma error from the greatest IRI value across the month.
- hmF2 of greater than 500 km or less than 200 km were removed as they are likely to be unphysical.

When a data point is removed from one parameter, the corresponding value is removed from all of the other time series.

Large errors in the data assimilation models can arise in two ways: undetected bad data can be assimilated, or good data can be assimilated badly. The first case includes the effects of undetected autoscaling errors. These can cause model errors directly and persist for a number of hours. Due to memory in the assimilative models, EDAM, GPSII and mNeQuick are all affected by these errors.

For EDAM, a small number of cases appear to be affected in the second way; i.e. reasonable virtual height profiles are assimilated, but result in grossly non-physical vertical electron density structures within the EDAM grid. As with the bad data case, the poor results can take a significant time to decay out of the assimilation, which leads to clumping of the outliers. It is not yet clear why such non-physical results occasionally arise in EDAM from the assimilation of good data. *McNamara et al.* [2013] chose to filter the EDAM results (for calculation of the RMS error) by excluding values that lay outside the two-sigma errors in the IRI values. However no such filtering has been applied to the results presented here.

It should also be noted that mNeQuick only produces results when valid ionosonde data is present, and as a result the mNeQuick model has far fewer data points than the others (2215 of a maximum of 2880 points, 77%).

It will be shown later, in Section 5.5, that both GITM and TIE-GCM sometimes produce unphysical results. These have been filtered out of the final statistics by removing hmF2 values greater than 500 km and less than 200 km. For GITM only, foF2 values less than 1 MHz have also been removed. It is not yet clear why GITM produces these unphysical electron density results for the nighttime foF2. Further discussion of the phenomena can be found in Section 5.5.

5.3. Observations

Table 3 shows the number of valid data points provided by the RSA ionosondes for September 2011, subject to the filtering described in Section 5.2. Figure 11 and Figure 12 show the time series plots of foF2 and hmF2 from the Grahamstown ionosonde.

Table 3. Number of valid data points produced by the ionosondes in the RSA for September 2011 subject to the filtering described in Section 5.2.

Station	Valid data points (out of 2880)	Percentage of valid points
Grahamstown (GR13L)	2492	86.5%
Hermanus (HE13N)	2758	95.8%
Louisvale (LV12P)	2793	97.0%
Madimbo (MU12K)	2428	84.3%

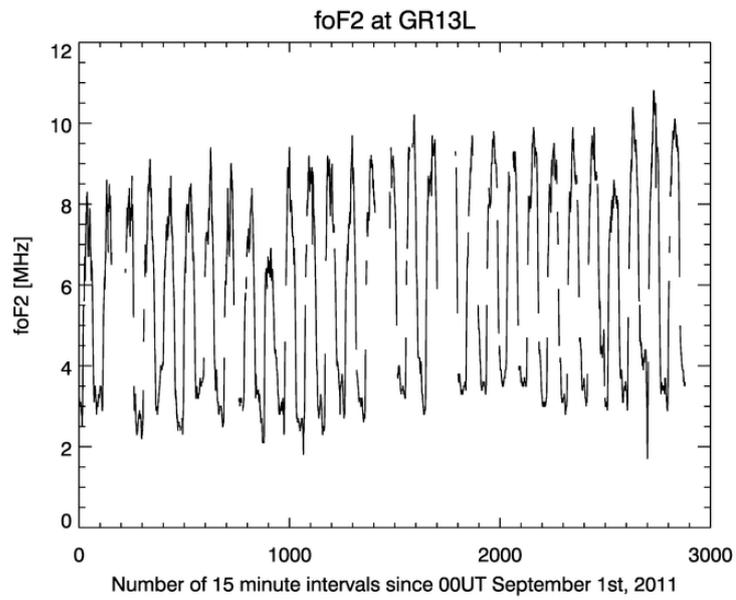


Figure 11. Reported foF2 time series from the Grahamstown ionosonde in the RSA, subject to the filtering described in Section 5.2.

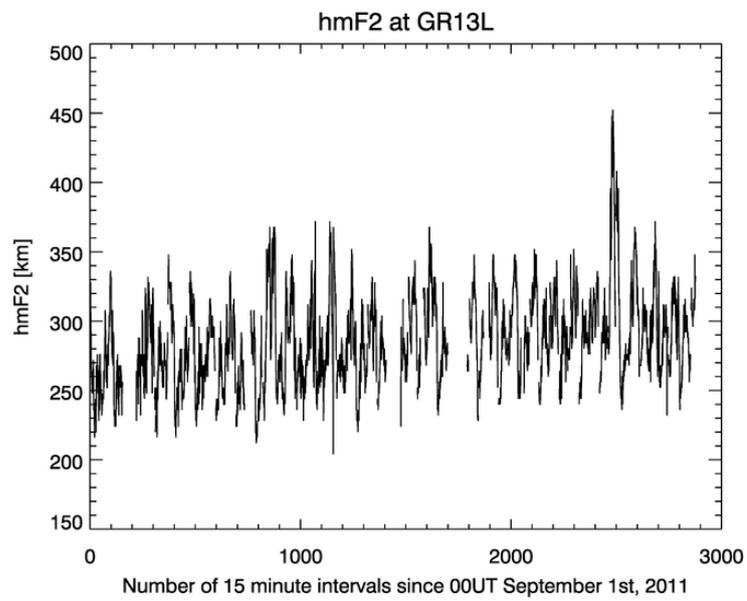


Figure 12. Reported hmF2 time series from the Grahamstown ionosonde in the RSA, subject to the filtering described in Section 5.2.

The foF2 values show a general increase over the course of the month. This is due at least partly to the corresponding increase in the F10.7 over the same time period (Figure 13). The trend in F10.7 can be fitted with a third order polynomial. The increase in F10.7 is believed to be the main cause for the non-stationarity of the foF2 time series.

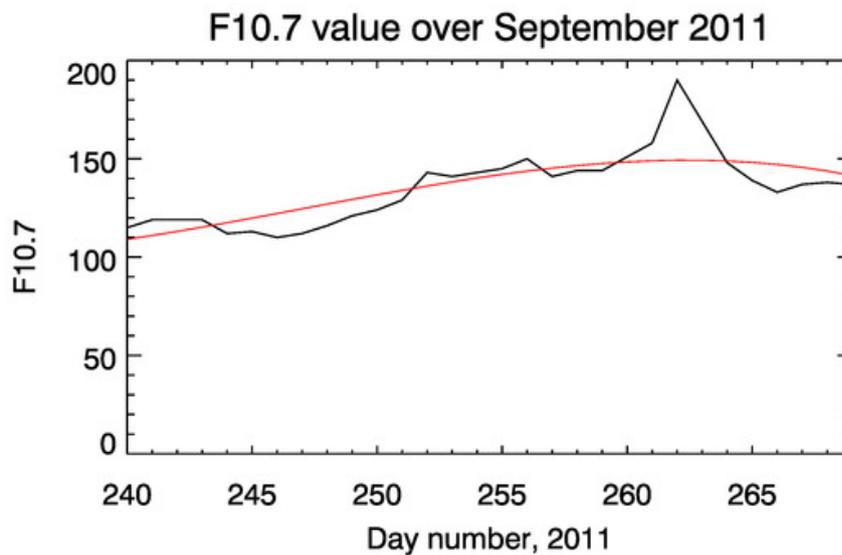


Figure 13. Black line is the observed F10.7 values over September 2011, the red line is a 3rd order polynomial fitted to the data.

The stationarity of the hmF2 and h(0.8foF2) time series is not dependent on F10.7, and instead depends on neutral winds and electric fields [Kelley, 2009]. The wind and electric field

variability is less severe during the test scenario, and, as such, the hmF2 time series is at least weakly stationary.

A time series is weakly stationary (or second-order stationary) if its mean is constant and the covariance depends only on the lag [Chatfield, 1991], i.e.

$$E[X(t)] = \mu, \quad (5.4)$$

$$\text{Cov}[X(t), X(t + \tau)] = \gamma(\tau). \quad (5.5)$$

Where $X(t)$ denotes a random variable at time t , μ is the mean and τ is the lag.

In order to apply the statistical techniques used in modified Taylor diagrams (Sections 4.2 and 4.3) the time series must be at least weakly stationary. Therefore a linearly scaled version of the fitted polynomial shown in Figure 13 was removed from the time series data to detrend it. The same trend was removed from each models' foF2 time series. The stationarity of all the time series was then tested using an Augmented Dickey-Fuller (ADF) test [Dickey and Fuller, 1979]. However before the test could be applied, the periodicity of each time series must be removed [Chatfield, 1991]. To accomplish this each model time series had the associated IRI time series removed from it, i.e. before applying the ADF test to the EDAM foF2 time series, the IRI foF2 time series was subtracted from the EDAM time series.

The critical value of the ADF test with a 95% confidence interval with a sample size greater than 500 is -3.96 [Enders, 2009]. That is, if a model time series has a critical value less than this the time series can be assumed to be stationary. The ADF critical values, before and after

the detrending are shown in Table 4. The hmF2 and h(0.8foF2) series are stationary without any detrending.

Table 4. Results from the ADF test for the model time series. For each foF2 time series, the value is greater than the critical value for the 95% confidence level (-3.96). However, after detrending, the time series is stationary.

	foF2		hmF2	h(0.8foF2)
	<i>Before detrending</i>	<i>After detrending</i>	<i>As is</i>	<i>As is</i>
EDAM	-2.52	-7.02	-14.50	-11.98
GPSII	-3.13	-8.79	-16.67	-13.36
mNeQuick	-2.96	-8.00	-16.18	-15.11
GITM	-3.90	-4.01	-20.00	-23.21
TIE-GCM	-2.03	-4.23	-5.37	-21.25

The autoscaled values of hmF2 are subject to the uncertainties of ionogram autoscaling and the conversion from the ionogram to the plasma frequency profile. Thus it can be difficult to quantify the level of uncertainty that exists in the ARTIST values of hmF2. Therefore the profile altitude at a plasma frequency of 80% of foF2 (64% of NmF2) is included as another validation parameter (h(0.8foF2)). h(0.8foF2) is expected to suffer from less natural and processing noise than hmF2, and has the advantage that it is a point on the F2 profile that is more relevant than hmF2 to HF radio propagation on oblique paths.

5.4. Comparison of IRI, EDAM, GPSII and mNeQuick

In this section, comparisons of an empirical model (IRI), and three data assimilation models (EDAM, GPSII and mNeQuick) are compared at each of the three truth sites (Hermanus, Louisvale and Madimbo).

5.4.1. Accuracy of Model Values at Hermanus

Figure 14 is the modified Taylor diagram using the Hermanus (HE13N) ionosonde for the truth data. It leads to a number of immediate conclusions: EDAM is slightly better than both GPSII and mNeQuick at specifying the values of foF2 since EDAM has a significantly stronger correlation, and a standard deviation closer to that of the observation. However GPSII shows no bias whereas EDAM and mNeQuick show a small negative bias. All three models show noticeable improvement over IRI.

The statistical significance between time series correlations can be found using the Fisher r-to-z transformation [Fisher, 1915, 1921]. For a large number of data points, as in this study, very small changes in correlation can be significant. A Z value of greater than or equal to $|1.96|$ when comparing two correlations implies they are significantly different at the 95% level [Kenny, 1987]. A table of the differences in correlation and corresponding Z values for the foF2 time series, using the Fisher r-to-z transformation, is shown in Table 5. Only the difference in correlation between mNeQuick and GPSII is not significant. Similar calculations can be done for the hmF2 and h(0.8foF2) time series.

Table 5. Z values from comparing the correlation coefficients for foF2 for each of the four models. Values greater than or equal to $|1.96|$ are significant at the 95% level.

	IRI		EDAM		GPSII		mNeQuick	
	Diff.	Corr.	Diff.	Corr.	Diff.	Corr.	Diff.	Corr.
IRI	-	-	-0.030	-29.41	-0.023	-18.40	-0.024	-19.18
EDAM	0.030	29.41	-	-	0.006	11.01	0.006	10.23
GPSII	0.023	18.40	-0.006	-11.01	-	-	-0.006	-0.186
mNeQuick	0.024	19.18	-0.006	-10.23	0.006	0.186	-	-

It can be further seen (Figure 14) that the standard deviation of model errors for mNeQuick in hmF2 is smaller than that of GPSII, EDAM and IRI (with the latter pair being very similar). The mNeQuick hmF2 value also has very little bias (~1 km) compared to IRI (~7 km), EDAM (~9.5 km) and GPSII (~6.5 km). Consequently mNeQuick shows considerable improvement over the other three models.

Using traditional analysis methods, such as the comparison of the root mean square error of EDAM and IRI (25 km and 23 km respectively) it can be erroneously concluded that EDAM is “defaulting” to its background values; as was reported in *McNamara et al.* [2013]. However it can be seen in Figure 14 that the EDAM hmF2 values offer an improvement in correlation over IRI, with a worse bias. Also, the modified Taylor diagram shows that EDAM tends to overestimate the range of hmF2 values, whilst IRI underestimates the range. This can be concluded since the hmF2 standard deviation is greater than unity for EDAM and less than

unity for IRI. Therefore the diagram shows that EDAM is not defaulting to the IRI values, and is manipulating them, even if it does not statistically improve them.

For the $h(0.8f_oF_2)$ parameter, Figure 14 reveals that EDAM slightly outperforms GPSII and mNeQuick, since it has a standard deviation very close to the observed value, and a smaller bias. The EDAM, GPSII and mNeQuick $h(0.8f_oF_2)$ values have similar correlation and all offer considerable improvement over IRI.

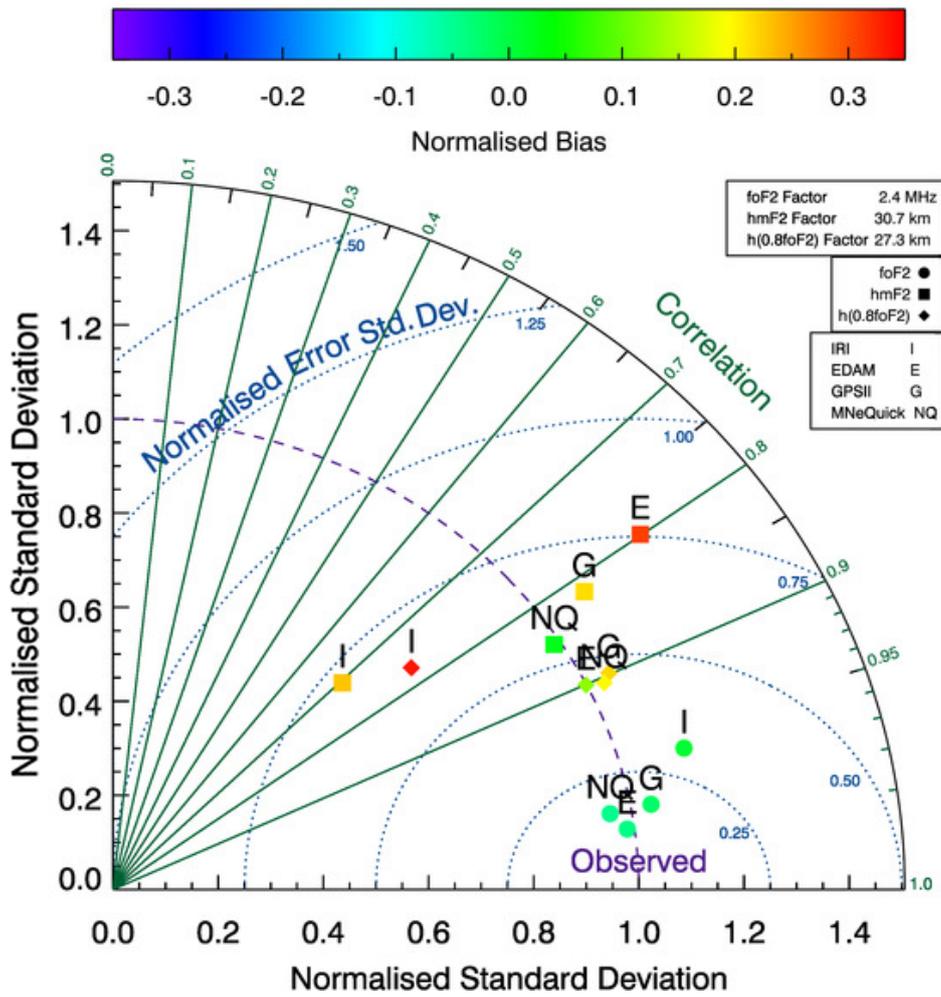


Figure 14. Modified Taylor diagram to show the model bias and standard deviation of errors. The azimuthal angle represents correlation, the radial distance the standard deviation and the semicircles centred at the 'Observed' marker the standard deviation of the errors. The colour scale shows the bias (mean of model minus mean of truth). To plot multiple parameters each quantity is normalized, the original values can be reformed using the corresponding 'factors' in the top right of the diagram. The diagram shows the second-order statistics for IRI, EDAM, GPSII and mNeQuick at Hermanus for the parameters foF2, hmF2 and h(0.8foF2).

McNamara et al. [2013] noted the large variation in the EDAM hmF2 results between day and night. Figure 15 shows the second-order statistics for EDAM with the day (08UT – 15UT),

night (20UT – 03UT) and dawn/dusk (03UT – 08UT and 15UT – 20UT) plotted separately. These times are chosen to correspond to the analysis performed in *McNamara et al.* [2013]. From the modified Taylor diagram it is apparent that EDAM performs much worse at day and night than at the combined dawn/dusk period; though the dawn/dusk period shows a large positive bias in $h(0.8f_oF_2)$ compared to the other times. EDAM performs worse at night due to the large relative errors in ionosonde measurements and worse during the day since the models are bound by the minimum expected errors set by travelling ionospheric disturbances (TIDs) [*Francis, 1975; Hunsucker, 1982; Hocke and Schlegel, 1999*]. On the other hand, at dusk and dawn, the modelling is dominated by the rapidly changing solar illumination conditions. For example, as the sun sets, and the radiance of X-rays and EUV drops considerably, the f_oF_2 values sharply fall. These rapidly changing conditions are accounted for accurately by the models.

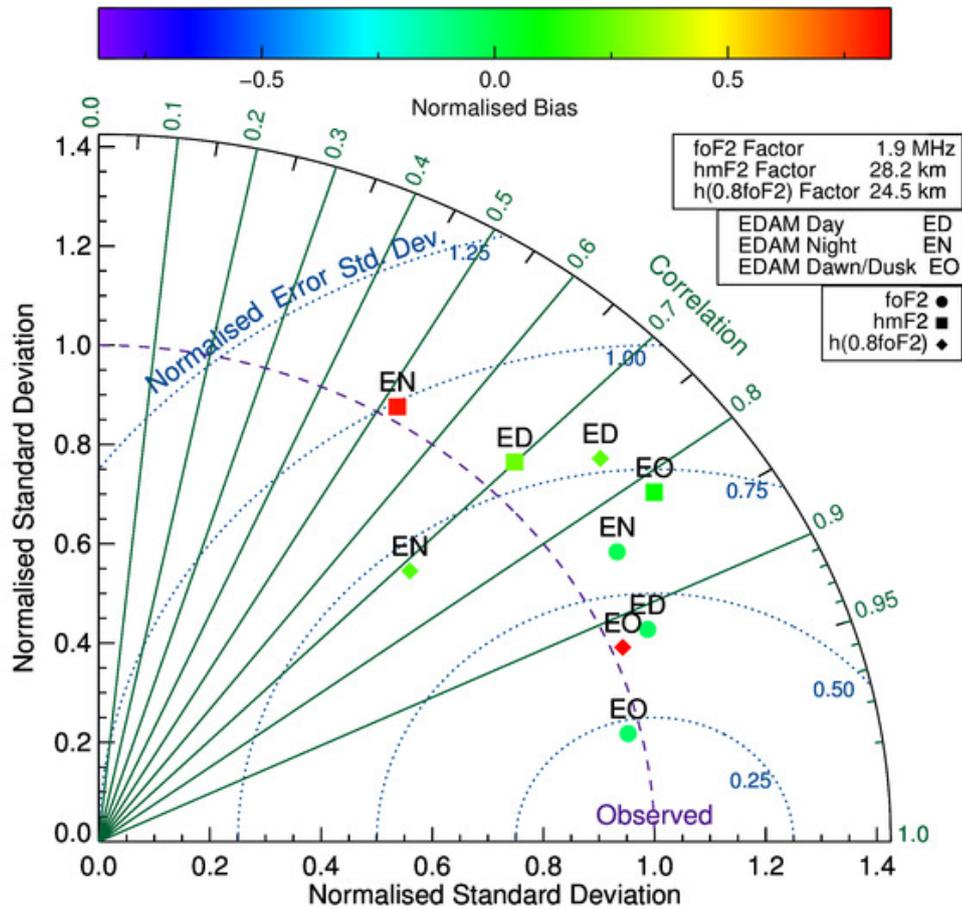


Figure 15. Modified Taylor diagram for Hermanus with second-order statistics split between day (08UT - 15UT), night (22UT - 5UT) and dawn/dusk (03UT - 08UT and 15UT - 20UT) for EDAM. The details of how to read the diagram are described in Figure 14.

Figure 16 again shows a modified Taylor diagram for the study period (September 2011); however, in this case, the second-order statistics for EDAM are plotted individually for each day of September. The over-plotted ovals show the grouping of each of the parameters (foF2, hmF2 and h(0.8foF2)). The foF2 daily values all have a similar correlation and standard deviation, as do the h(0.8foF2) values, albeit with slightly more variation. However the hmF2 values have a much larger spread, in standard deviation, correlation and bias. This accounts for

the poor hmF2 results EDAM exhibit in the whole month analysis (Figure 14). *McNamara et al.* [2013] commented that the poor EDAM results were likely to be due to large outliers. However Figure 16 shows that outliers are not skewing the result and that EDAM performs relatively poorly throughout the month. A similar modified Taylor diagram can be plotted for GPSII (Figure 17). This contains the same over-plotted ovals as in Figure 16, so a direct comparison between the EDAM and GPSII parameters can be made. As expected, the foF2 and h(0.8foF2) results are similar. However GPSII's hmF2 values have much less variability than EDAM, which results in GPSII's considerable improvement in hmF2 over EDAM at Hermanus. Figure 18 shows the individual days for mNeQuick which shows a close grouping of each parameter; though there is an outlier for h(0.8foF2) (in terms of correlation and standard deviation, corresponding to September 6th) and two outliers (in terms of bias, corresponding to 19th & 20th September). The individual day modified Taylor diagram for IRI (Figure 19) provides an interesting insight into the IRI results. It can be seen that the variability of the heights (both hmF2 and h(0.8foF2)) are always underestimated by IRI whereas the variability of foF2 is, in all but three cases, overestimated. This is because IRI overestimates the nighttime hmF2 values, and underestimates the daytime hmF2. It is not currently clear why IRI does this, or if it is generally true, and further investigation is required. Figure 16 and Figure 19 also provide insight into the effectiveness of the EDAM assimilation. Since IRI is the background model to EDAM we would expect EDAM's use of data to improve the results in specifying the ionosphere. It can be seen from comparing Figure 16 and Figure 19 that this is true for foF2 and h(0.8foF2). However, for hmF2, EDAM is "pushing" the results in the right direction, but by too much. Hence the values go from less than unity to greater than unity (in standard deviation) with no obvious improvement in correlation or bias.

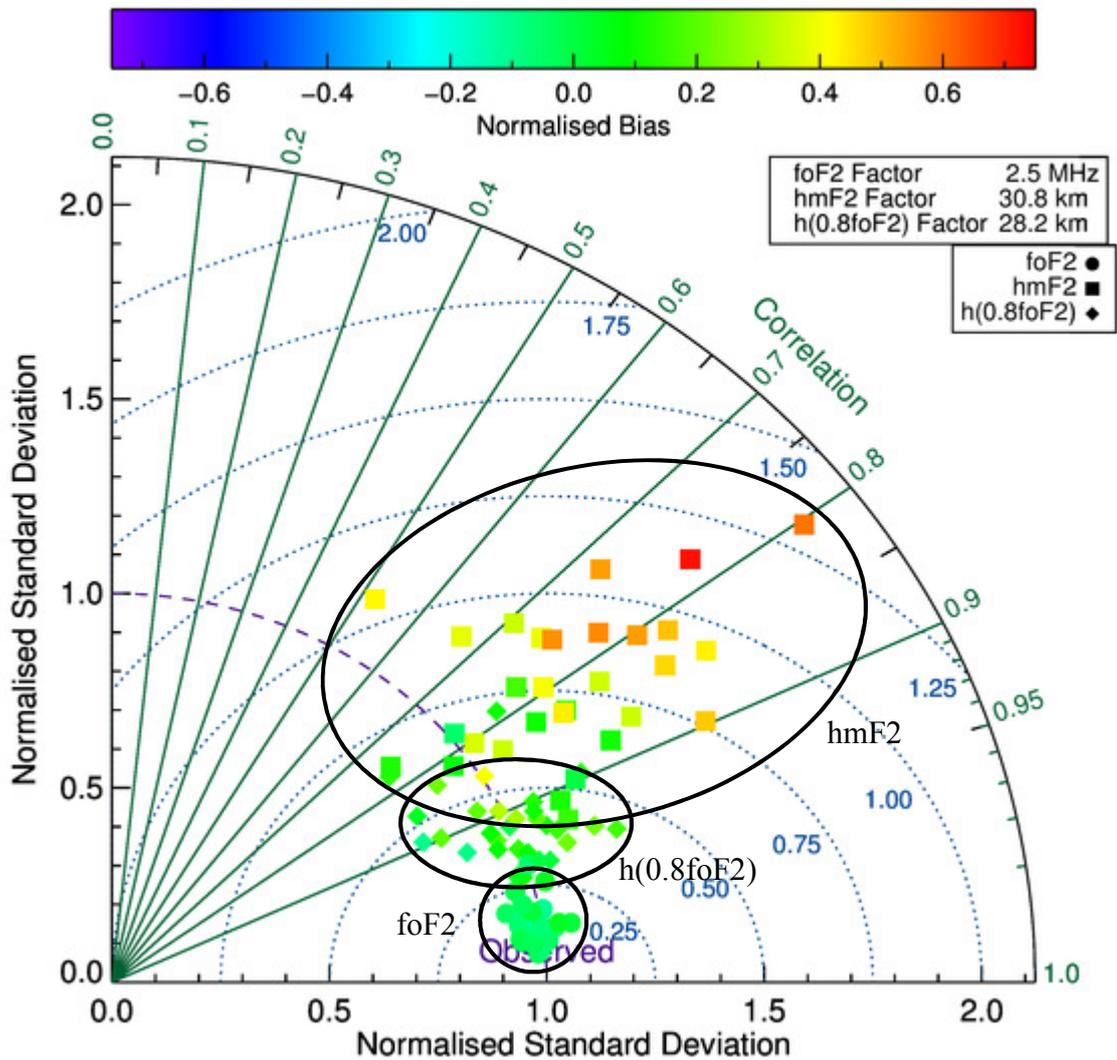


Figure 16. Modified Taylor diagram of the second order statistics for EDAM at Hermanus. Each day of the test study (September 2011) is plotted individually. The details of how to read the diagram are described in Figure 14.

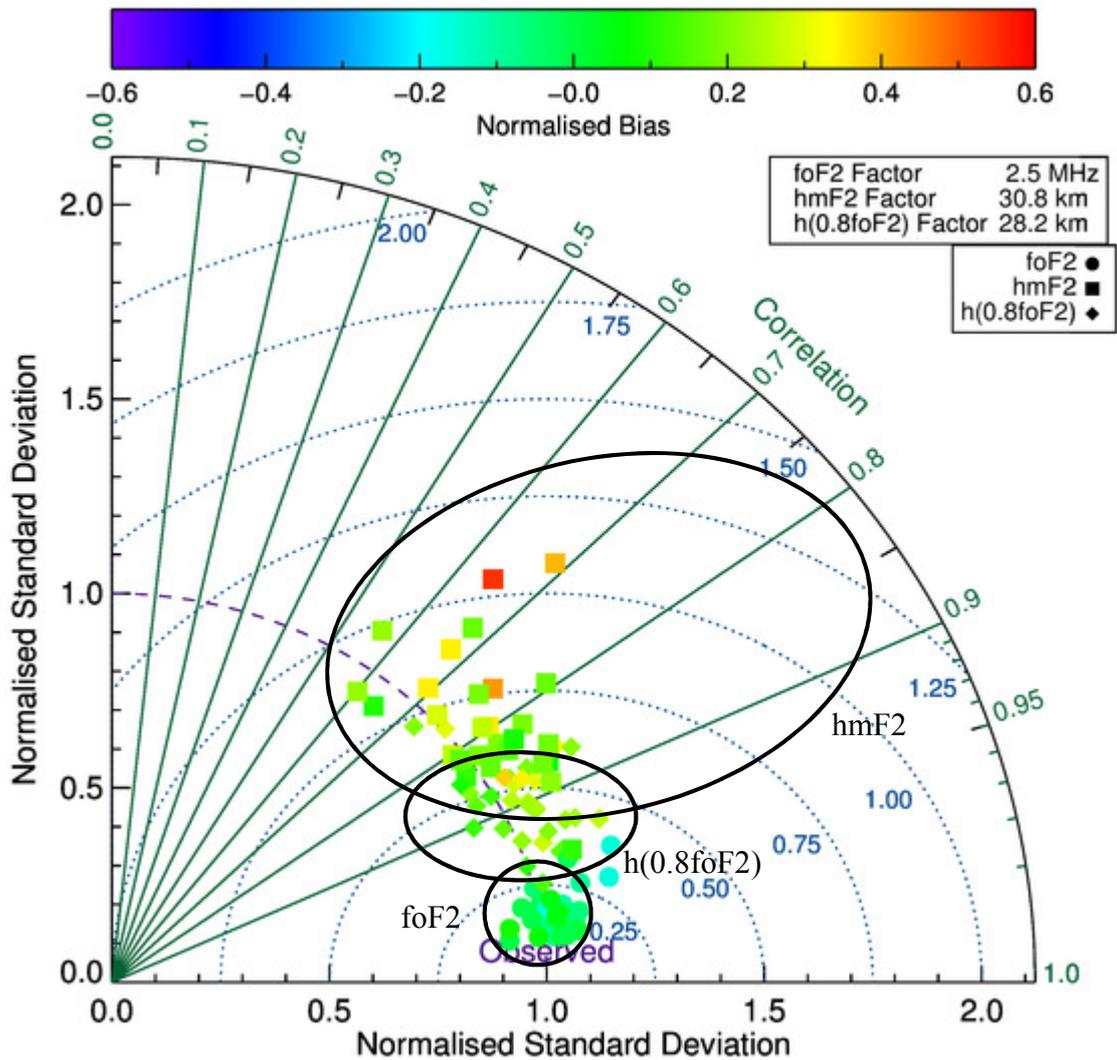


Figure 17. Modified Taylor diagram of the second order statistics for GPSII at Hermanus. Each day of the test study (September 2011) is plotted individually. The over plotted ovals are the same as in Figure 16. The details of how to read the diagram are described in Figure 14.

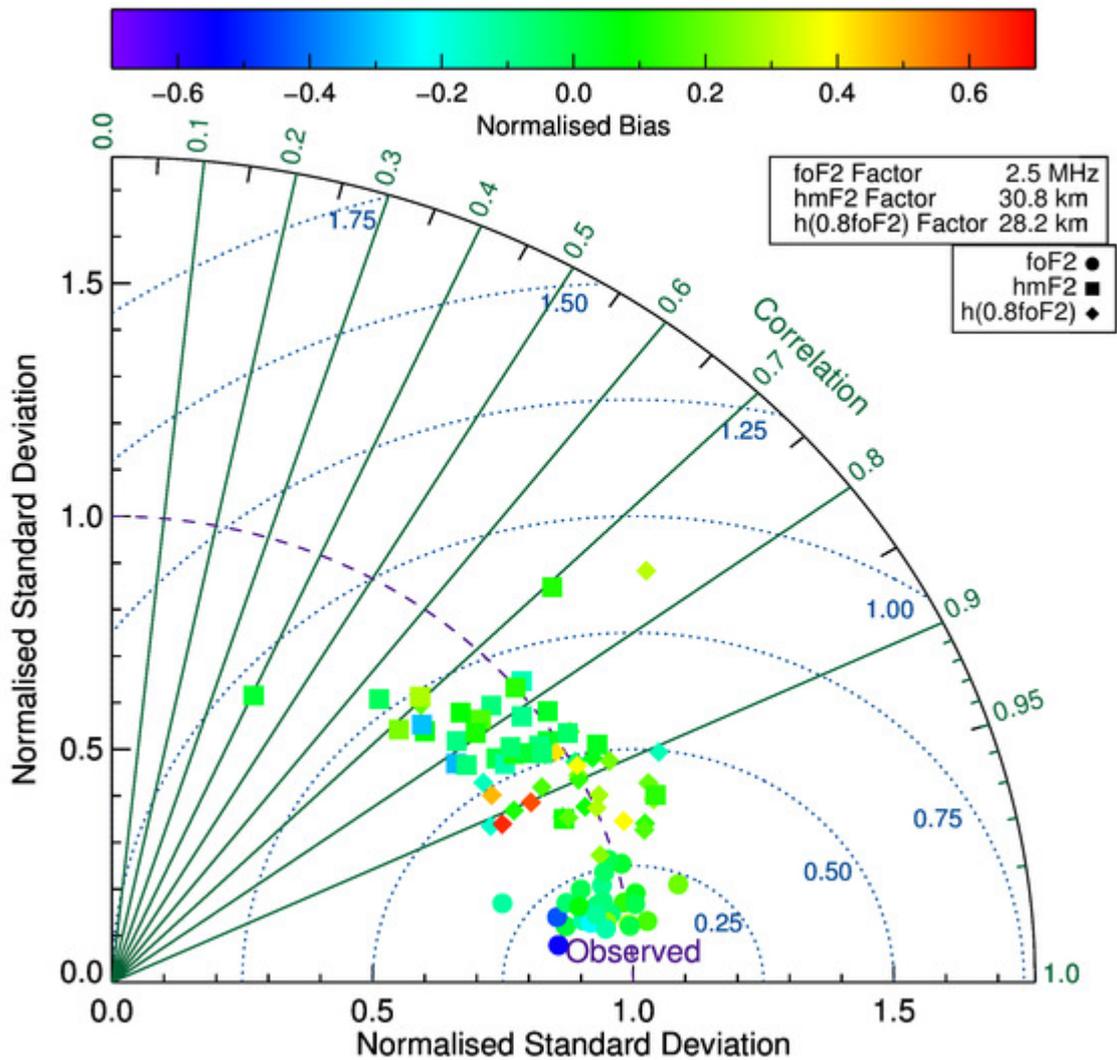


Figure 18. Modified Taylor diagram of the second order statistics for mNeQuick at Hermanus. Each day of the test study (September 2011) is plotted individually. The details of how to read the diagram are described in Figure 14.

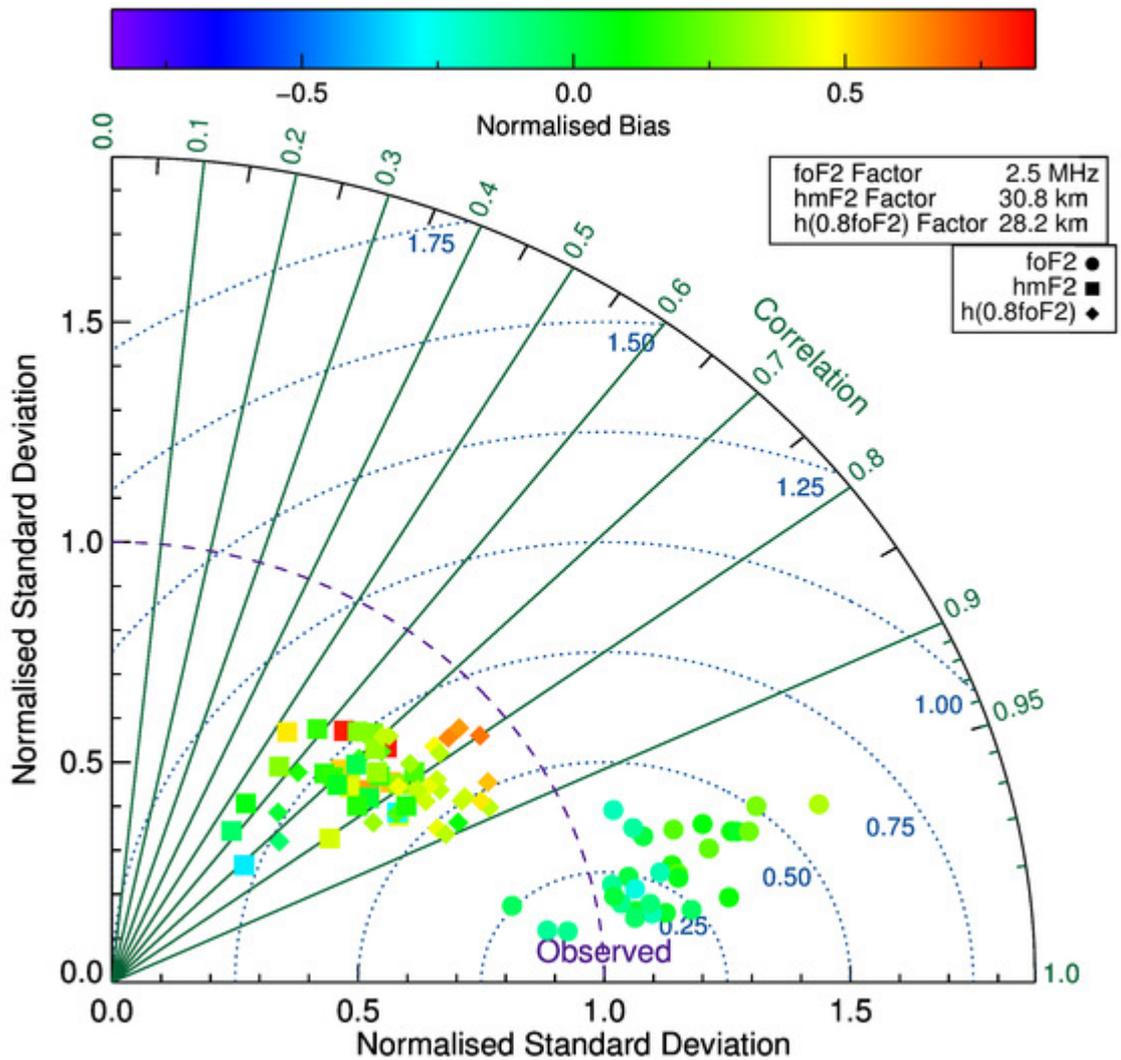


Figure 19. Modified Taylor diagram of the second order statistics for IRI at Hermanus. Each day of the test study (September 2011) is plotted individually. The details of how to read the diagram are described in Figure 14.

5.4.2. Accuracy of Model Values at Louisvale

Figure 20 shows the modified Taylor diagram for the Louisvale (LV12P) station. EDAM, GPSII and mNeQuick are very similar at specifying the foF2 values both in terms of standard deviation and correlation, although all are worse than at Hermanus.

This decrease in performance is to be expected since approximate correlation lengths depend on whether the stations are north-south of each other or east-west (Table 6 [McNamara, 2009]). The Louisvale station is ~735 km north-west of Grahamstown, compared to Hermanus which is ~700 km west.

Table 6. Suggested correlation lengths for use with global models. Reproduced from Table 3 in McNamara [2009].

Solar Activity	North-South (km)	East-West (km)
High	1000	1500
Low	700	1000

Analysis of foF2 in the modified Taylor diagram (Figure 20) shows that although the performances are statistically similar, the models respond to the data differently: the standard deviation for GPSII and mNeQuick is greater than the observation whereas for EDAM it is lower. This implies that the EDAM range of values for foF2 is smaller than that observed, whilst GPSII and mNeQuick have a greater range.

For hmF2, EDAM shows a considerable improvement over its results for Hermanus in terms of standard deviation, albeit with a slightly worse correlation. EDAM shows a better standard deviation than both GPSII and mNeQuick (though with a greater bias) whereas GPSII and mNeQuick have a stronger correlation. The mNeQuick model is the only one of the three to show no significant bias. The investigation from *McNamara et al.* [2013] concluded that the EDAM errors did not show any improvement over IRI; this conclusion was based solely on analysis of the RMS errors. However it can be seen from the modified Taylor diagram that EDAM does, in fact, offer an improved correlation and standard deviation over IRI.

Finally, Figure 20 shows a reversal in performance in the $h(0.8f_oF_2)$ parameter as compared to Hermanus. mNeQuick and GPSII both outperform EDAM; considerably in terms of standard deviation and slightly with regards to correlation. However, all three again show an improvement over IRI.

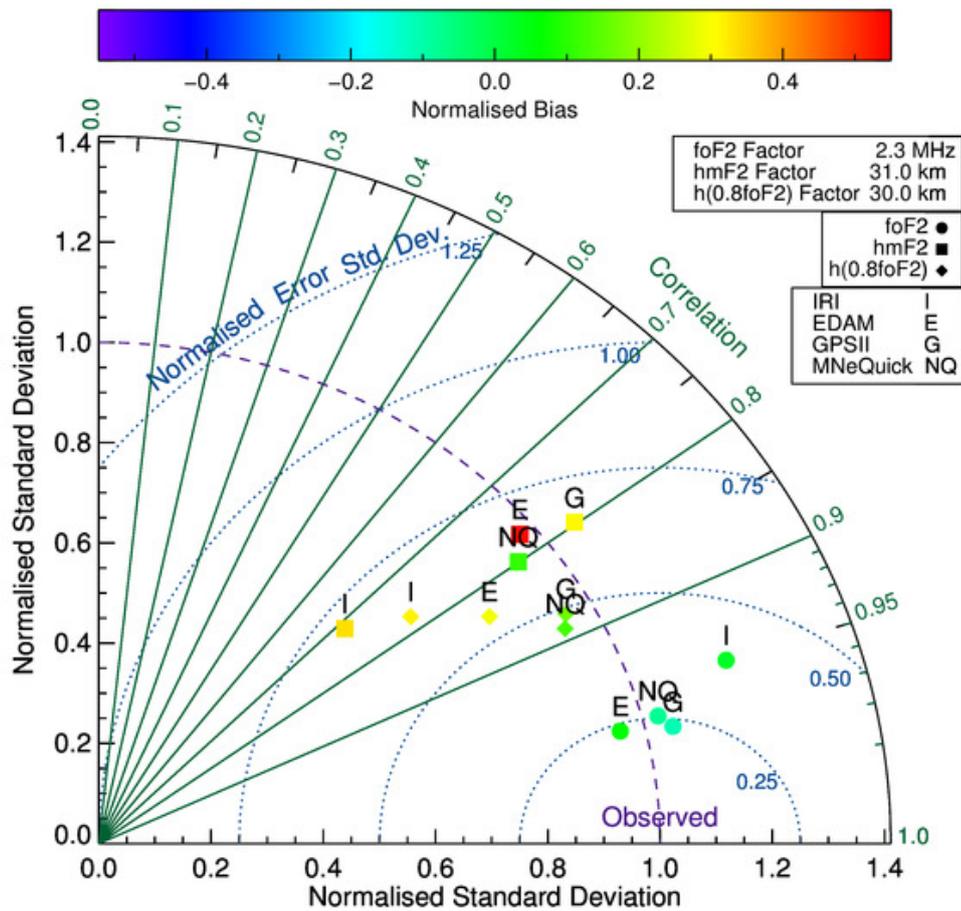


Figure 20. Modified Taylor diagram for IRI, EDAM, GPSII and mNeQuick at Louisvale for the parameters foF2, hmF2 and h(0.8foF2). The details of how to read the diagram are described in Figure 14.

5.4.3. Accuracy of Model Values at Madimbo

The Madimbo station is just under 1300 km to the north of the assimilated Grahamstown station so it was expected that this station would show the worst results (see discussion in Section 5.4.2). The modified Taylor diagram for this station is shown in Figure 21.

At this station, GPSII outperforms EDAM and IRI in all parameters, except the error standard deviation, and mNeQuick in foF2 and hmF2. mNeQuick shows a negative bias in all parameters. In foF2, GPSII is clearly the best, with mNeQuick second, whilst EDAM offers no improvement in the RMS error over IRI. However EDAM does show a significant improvement, at the 95% confidence level, in correlation over IRI (as discussed at the start of Chapter 5). It can be seen from the modified Taylor diagram that EDAM has a smaller (than unity) normalised standard deviation, whilst IRI's is greater than unity. This again shows that EDAM is not defaulting to the IRI values, and is still modifying its background model.

In hmF2 EDAM shows a very small improvement over IRI, with GPSII and mNeQuick clearly outperforming the others, except for the error standard deviation. Finally, for the h(0.8foF2) parameter, EDAM and IRI are almost identical (the points are inseparable), and the mNeQuick and GPSII parameters are also very similar.

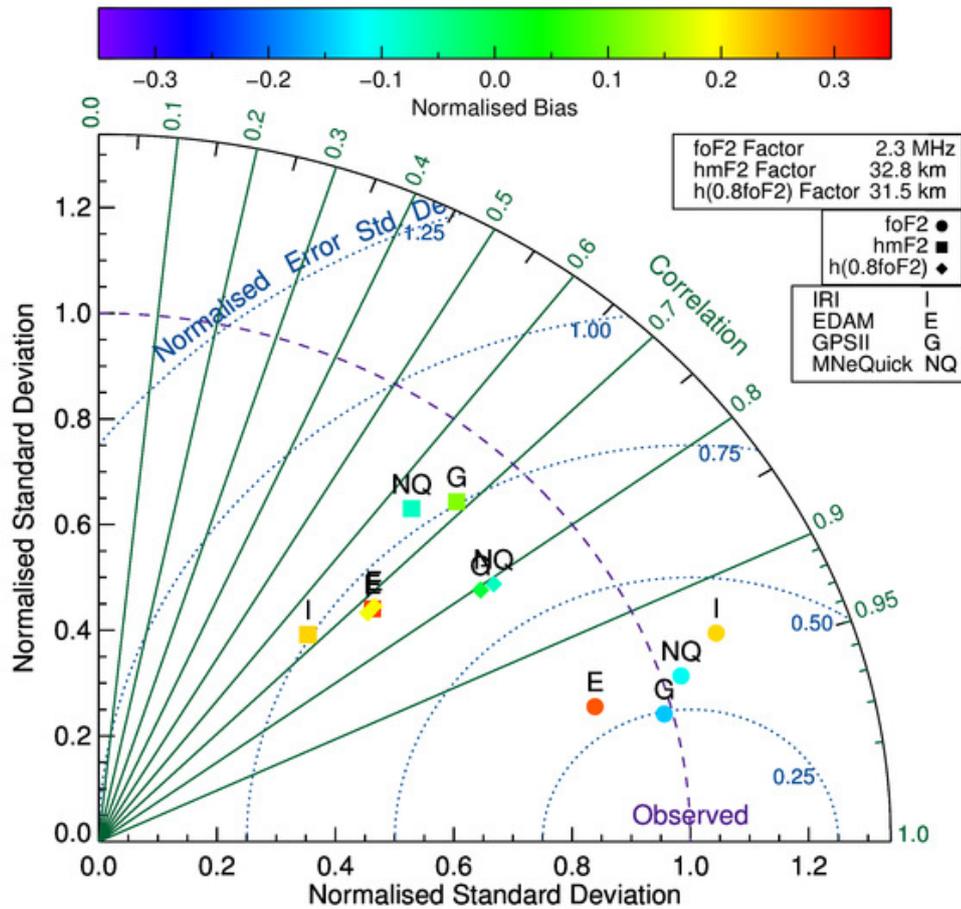


Figure 21. Modified Taylor diagram for IRI, EDAM, GPSII and mNeQuick at Madimbo for the parameters $foF2$, $hmF2$ and $h(0.8foF2)$. The details of how to read the diagram are described in Figure 14.

5.5. Comparisons Including TIE-GCM and GITM

The physics-based models tested in this chapter are not specifically ionospheric models. The models are atmospheric models, and include neutral and ion species, as well as temperature and velocity. Electron density values are calculated simply as the sum of the ion densities. Unlike IRI they are not calibrated with ionospheric measurements (i.e. the CCIR coefficients). The model comparison is shown in Figure 22.

In foF2, TIE-GCM and GITM both have similar correlations with the observations. However TIE-GCM gives a standard deviation of less than unity, implying the model tends to underestimate the range of values for foF2, whilst GITM is greater than unity, i.e. it overestimates the range. TIE-GCM has a smaller standard deviation of errors (model minus observation, ~1.25 MHz), whereas GITM's standard deviation of errors is ~2 MHz. GITM also has a negative bias in foF2.

For hmF2 the model performance is more spread. IRI and TIE-GCM have similar RMS errors when specifying hmF2; however it can be seen from the modified Taylor diagram that they are different. The standard deviation of errors are very similar however IRI has no noticeable bias whereas TIE-GCM has a large positive bias (~30.8 km). In terms of correlation IRI and TIE-GCM are similar. Finally, GITM performs poorest in terms of correlation with the observation and standard deviation of errors.

For h(0.8foF2), TIE-GCM produces a standard deviations close to the observation, closer than both IRI and GITM. The TIE-GCM result has a correlation of ~0.62 and shows a positive bias (in line with its hmF2 results). GITM however shows very little correlation with the observation (~0.47) and a negative bias. IRI gives a standard deviation less than unity but has a better correlation and mean square error than both TIE-GCM and GITM.

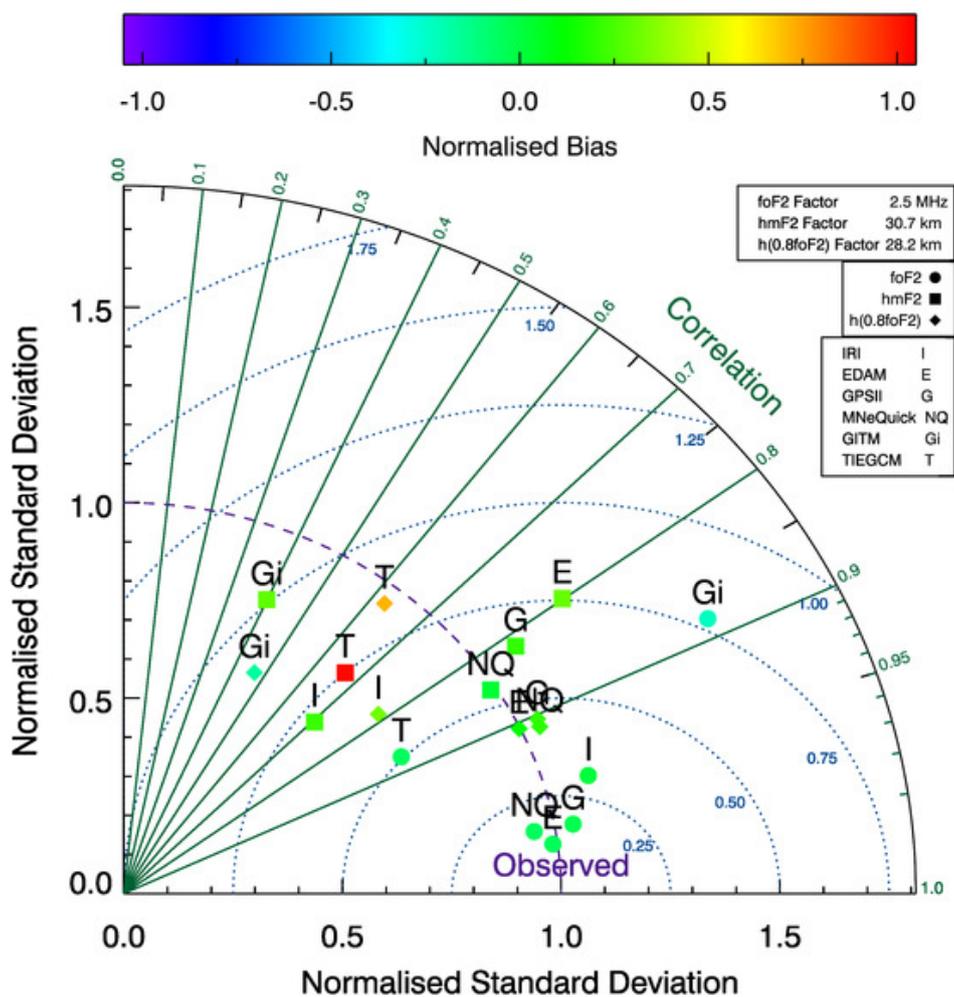


Figure 22. Modified Taylor diagram for IRI, EDAM, GPSII, mNeQuick, GITM and TIE-GCM at Hermanus for the parameters foF2, hmF2 and h(0.8foF2). The details of how to read the diagram are described in Figure 14.

The tested physics models both perform worse than the empirical and data assimilative models tested in the previous sections. However the models still perform fairly well, especially TIE-GCM. This indicates that such models could be used as a background model in data assimilation schemes for ionospheric forecasts in the future.

GITM regularly produces nighttime foF2 values of less than 1 MHz. Therefore, as described in Section 5.2, the results have had a filter applied to them to remove unphysical results. However even with these values removed the GITM results still look unphysical at night (Figure 23). This, at least in part, explains GITM's poor performance as compared to the other models.

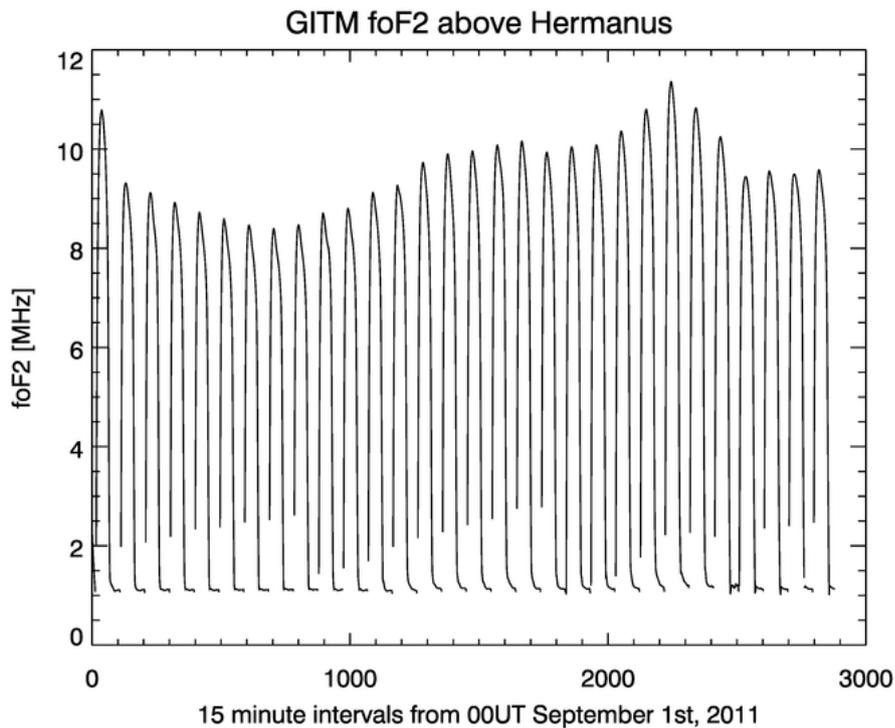


Figure 23. Time series of GITM's foF2 values across the month of September 2011. The nighttime values have unphysical characteristics.

5.6. Comparisons Including Model Combinations

As introduced in Section 3.4, expecting model performance to improve by combining results is based on the assumption that errors should cancel. Simple combinations are discussed here. More formal multi-model ensembles will be discussed in Chapter 6.

A naïve approach to model combination is to simply take two time series (i.e. the EDAM and GPSII foF2 values above Hermanus) and take the mean of the two at each time step. The new combination time series can then be compared to the observations. The results of combining four model time series (IRI, EDAM, GPSII and mNeQuick) for foF2, hmF2 and h(0.8foF2) are shown in Figure 24. The combined result is labelled ‘C’.

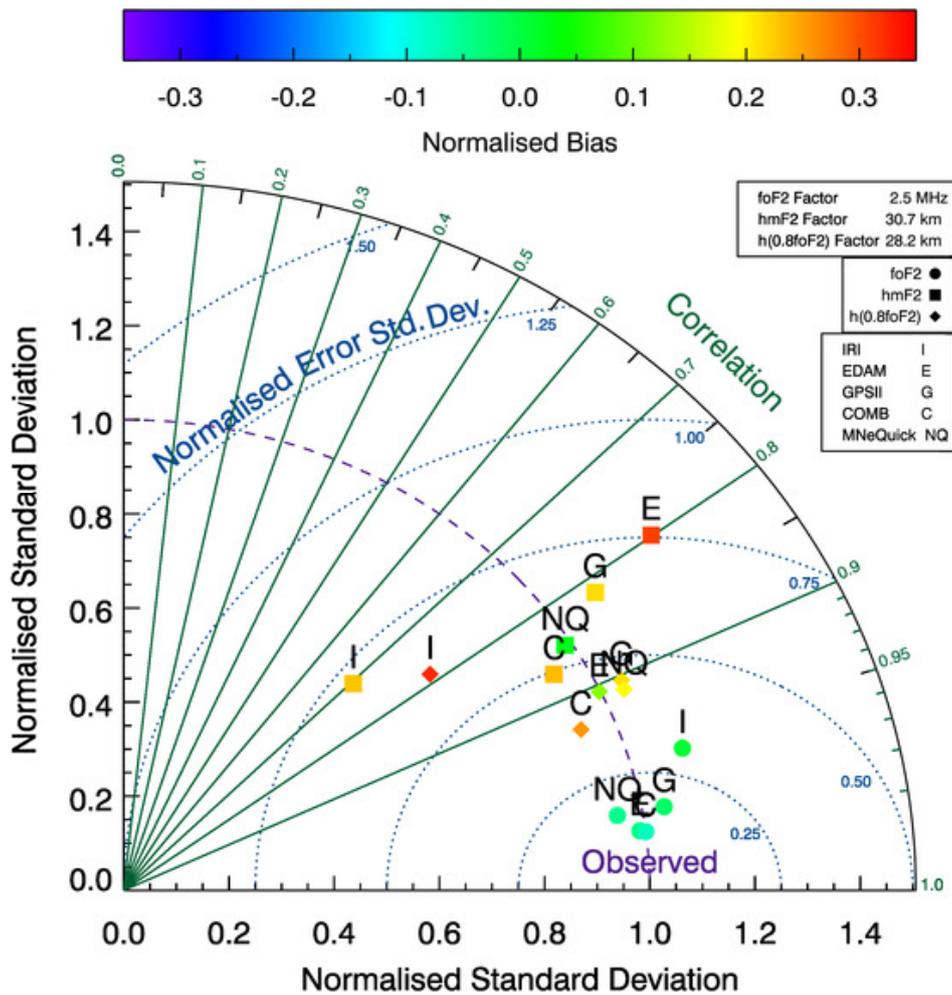


Figure 24. Modified Taylor diagram for IRI, EDAM, GPSII, mNeQuick and the average combination of their parameter time series at Hermanus for the parameters foF2, hmF2 and h(0.8foF2). The details of how to read the diagram are described in Figure 14.

It is clear that combining the parameters does no harm to the results. Across all three of the tested parameters the combination offers a statistical improvement in correlation compared to the individual models, except for EDAM in foF2. In terms of bias, this approach leads to the combination bias being the average of the contributing model bias. Consequently, the combination performs worse than some of the models, and better than others. In foF2 the combination performs similarly to EDAM, the best of the tested models. In both hmF2 and h(0.8foF2) the combination has a slight positive bias with a standard deviation of less than unity. This is due to IRIs small standard deviation as compared to the observations. The model parameters which have made up the combination have a large spread in hmF2 and h(0.8foF2). Thus the improvement in correlation, and providing standard deviations close to unity (much closer than some of the constituent models) shows the technique provides valuable results.

However, for models to be useful to operational RF systems, the whole electron density profile is often required, rather than just the three parameters tested so far. Therefore, the EDAM and GPSII full electron density profiles across the month have been combined to produce a new profile; i.e. at each altitude (90 – 400 km, in 4 km steps) the electron density values from EDAM and GPSII have been averaged (Figure 25).

The modified Taylor diagram of the results (Figure 26) shows that for this type of combination the correlation of all parameters' are improved, again except for EDAM's foF2. For foF2, the combination is comparable to the EDAM results, but with a standard deviation closer to the observation. In hmF2, the combination is considerably better than EDAM, and offers an improvement in correlation over GPSII. Finally for h(0.8foF2) the combination improves on GPSII the most, but also offers an improvement over EDAM.

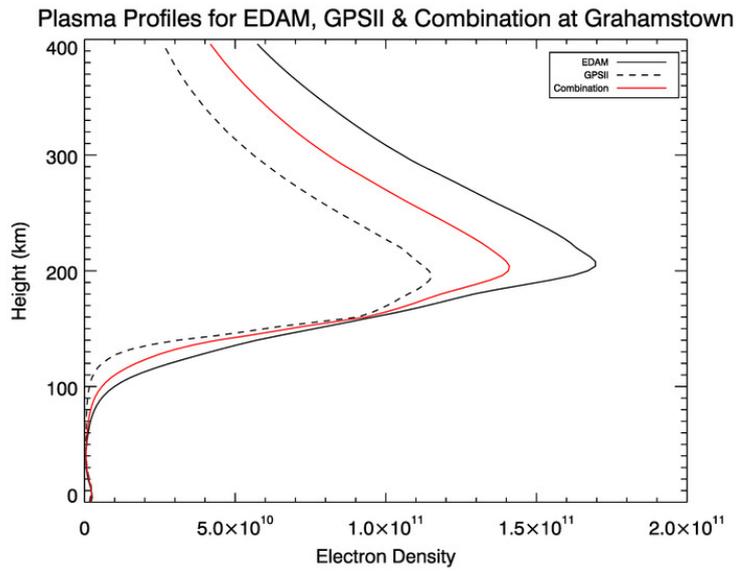


Figure 25. Example electron density height profile, black solid line is from EDAM and dashed line from GPSII. The red solid line is the average of the two.

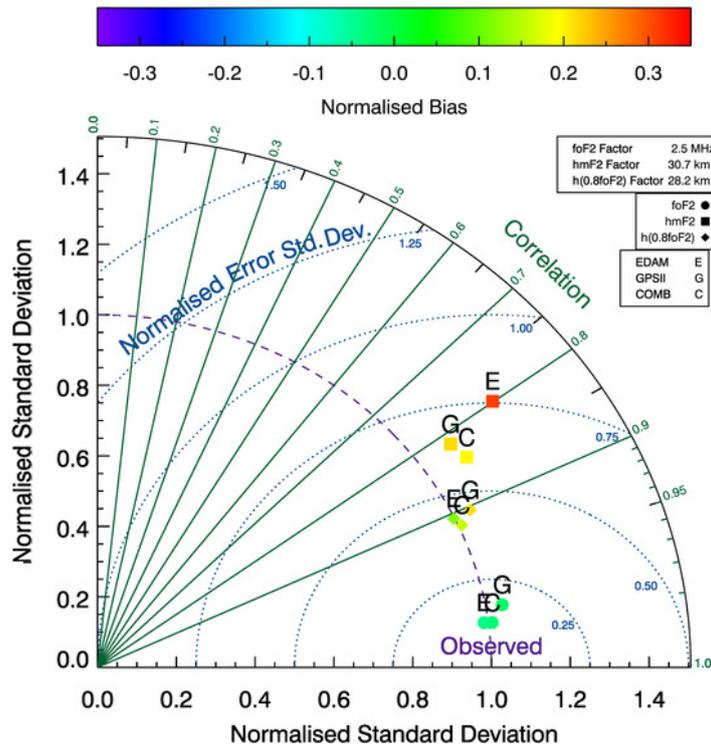


Figure 26. Modified Taylor diagram for EDAM, GPSII and the average combination of their height profiles at Hermanus for the parameters foF2, hmF2 and h(0.8foF2). The details of how to read the diagram are described in Figure 14.

5.7. Conclusions

Before statistical analysis can be performed, and models compared, care should be taken to ascertain the stationarity of the time series in question. This can be accomplished by using the Augmented Dickey-Fuller (ADF) test. It has been found for this test scenario that the hmF2 and h(0.8foF2) time series do not require detrending, however the foF2 time series does.

Empirical, physics-based and data assimilation models have been tested for a one month test scenario. The three assimilative models, EDAM, GPSII and mNeQuick all show improvement over IRI. Even though EDAM does not statistically improve upon IRIs hmF2 values, GPSII and mNeQuick do. It has been shown that the poor EDAM hmF2 results are not due to any single day outliers, and that the dawn/dusk time period is better modelled than day or night. The expected decrease in model performance, as a factor of distance from the assimilated ionosonde site, has also been shown.

TIE-GCM has been shown to perform only slightly worse than IRI, and comparable for hmF2. Since IRI is currently used as the background model for a number of data assimilation models (e.g. EDAM and GPSII), these results suggest that TIE-GCM could be used as a physics-based background model for future data assimilation models. It is advantageous to use a physics-based background model so that the electron densities can be more accurately propagated forward in time, providing forecasts. GITM, on the other hand, has performed worse than TIE-GCM, at least partly due to its unphysical results.

Finally, combining both model parameters and profiles has shown to make a statistical improvement in the model results. A more sophisticated method of model combination is that

of multi-model ensembles (MMEs). This approach uses a combination of model results to reinitialise a physical model and is the subject of Chapter 6.

6. MULTI-MODEL ENSEMBLES

The whole is more than the sum of its parts.

- Aristotle, *Metaphysica*

As shown in Sections 5.5 and 5.6, combining model results can lead to an improved performance. A more sophisticated approach than the simple model combination already discussed is that of multi-model ensembles (MMEs). By combining the number density of individual ions, neutrals, temperatures and velocities, the model errors may be reduced.

There are many approaches to constructing an MME [*Krishnamurti, 1999; Rajagopalan et al., 2002; Barnston et al., 2003*]; two are described in this chapter. The MMEs are compared to the individual model runs and also used as the initial conditions in the physics-based model TIE-GCM. The MME should have lower number density errors, and thus initializing a model with the improved conditions should result in improved forecasting capabilities.

The work presented in this chapter is an expanded version of *Elvidge et al.*[2013].

6.1. Atmospheric Neutral Densities

NASA predicts that, by 2030, orbital collisions could become frequent enough to cause a cascade, known as the Kessler Syndrome [*Kessler et al., 2010*], with the potential to prevent the use of low Earth orbit (LEO) [*Koller, 2012*]. One way to prevent the Kessler Syndrome is to more accurately predict orbital trajectories to better plan satellite collision avoidance manoeuvres. A key component in orbital trajectory predictions is an accurate description of the upper atmospheric environment, in particular the ionosphere-thermosphere, since drag due to

atmospheric density is one of the main forces that affect the orbit of satellites and space debris. The neutral air density from 200 to 1000 km altitude can change by 80% diurnally as well as by two to three orders of magnitude during geomagnetic storms; sometimes in just a few hours [Sutton *et al.*, 2005]. The upper atmosphere forecast models currently in use for orbit prediction are empirical [NASA *Space Vehicle Criteria (Environment)*, 1973], and include NRLMSISE-00, the Jacchia Reference Atmosphere [Jacchia, 1977] and the NASA/MSFC Global Reference Atmospheric Model-1999 Version [Justus and Johnson, 1999]. They are finely tuned, but when applied to satellite orbit forecasts they can result in large uncertainties in the orbital parameters (positional errors on the order of kilometres after a day [Vallado and Finkleman, 2008; McLaughlin *et al.*, 2011]).

Techniques exist to improve the predictability of the upper atmospheric environment, for example bootstrapping [Sunil Rao, 2000]. However MMEs are used here to enhance the prediction of the full ionosphere-thermosphere system. The main objective is to minimize the effects of model errors and bias, and improve the prediction of the physical phenomena using the ensemble.

6.2. Test Scenario

The time period for this neutral density study was from August 28th 2009 to September 1st 2009. This was during solar minimum, where it is thought there is the largest difference in reported densities between the models [Godinez, 2013]. At solar minimum the impact of the solar input parameters on the models is relatively small, and thus other internal and external dynamics dominate the evolution of the ionosphere-thermosphere densities. This particular time period was chosen since it included a geomagnetic storm which took place on August 30th. The

Ap index reached a high of 67 between 15UT and 18UT August 30th, whilst staying below 10 at other times. The F10.7 showed little variability throughout the whole test period (Figure 27).

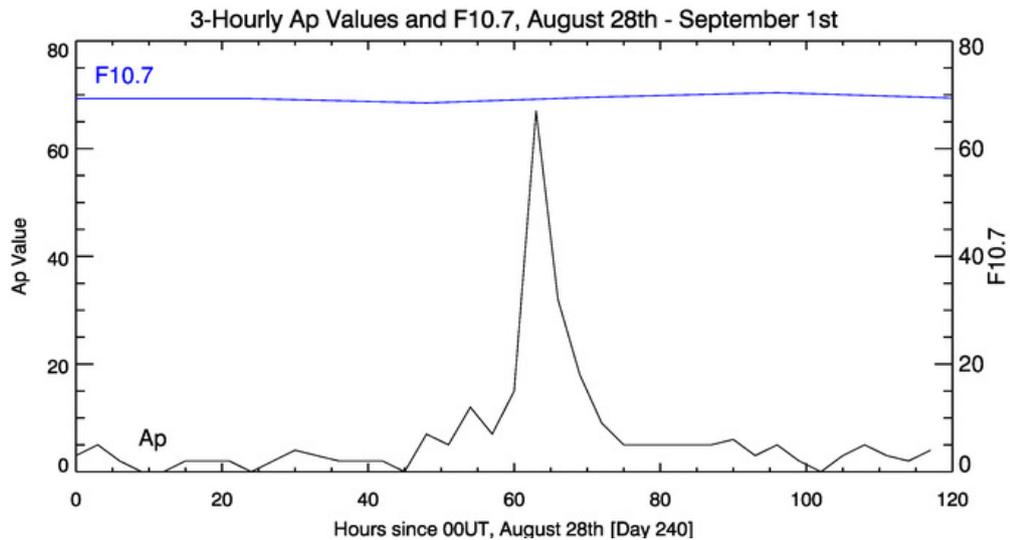


Figure 27. Ap and F10.7 for the neutral density MME test study, August 28th – September 1st 2009. The spike in Ap is due to a geomagnetic storm.

6.3. Models and Observations

The models tested in this scenario are NRLMSISE-00 (described in Section 3.1.3 [Picone *et al.*, 2002]), TIE-GCM (Section 3.2.1 [Roble *et al.*, 1988; Richmond *et al.*, 1992]) and GITM (Section 3.2.2 [Ridley *et al.*, 2006]). In this test the models were run with a 30 minute time step.

The models have a wide variety of outputs as shown in Table 7. It is important to note that the models do not have every output in common. This plays an important part when constructing an MME (discussed in Section 6.7.1.). The inputs for each model, described in Chapter 3, are presented in Table 8.

Table 7. NRLMSISE-00, TIE-GCM and GITM model outputs. mmr is the mass mixing ratio.

NRLMSISE-00	TIE-GCM	GITM
He (cm ⁻³)		He (m ⁻³)
O (cm ⁻³)	O (mmr)	O (m ⁻³)
O ₂ (cm ⁻³)	O ₂ (mmr)	O ₂ (m ⁻³)
N (cm ⁻³)	N (mmr)	N (m ⁻³)
N ₂ (cm ⁻³)	N ₂ (mmr)	N ₂ (m ⁻³)
Ar (cm ⁻³)		
H (cm ⁻³)		H (m ⁻³)
	NO (mmr)	NO (m ⁻³)
	O ⁺ (cm ⁻³)	O ⁺ (m ⁻³)
	O ₂ ⁺ (cm ⁻³)	O ₂ ⁺ (m ⁻³)
		N ⁺ (m ⁻³)
		N ₂ ⁺ (m ⁻³)
	NO ⁺ (cm ⁻³)	NO ⁺ (m ⁻³)
	N _e (cm ⁻³)	N _e (m ⁻³)
Neutral temp. (K)	Neutral temp. (K)	Neutral temp. (K)
	Ion temp. (K)	Ion temp. (K)
	Electron temp. (K)	Electron temp. (K)
	Neutral meridional wind (cms ⁻¹)	Neutral velocity (east) (ms ⁻¹)
	Neutral zonal wind (cms ⁻¹)	Neutral velocity (north) (ms ⁻¹)
	Neutral vertical wind (cms ⁻¹)	Neutral velocity (up) (ms ⁻¹)
		Ion velocity (east) (ms ⁻¹)
		Ion velocity (north) (ms ⁻¹)
		Ion velocity (up) (ms ⁻¹)
		O velocity (up) (ms ⁻¹)
		O ₂ velocity (up) (ms ⁻¹)
		N velocity (up) (ms ⁻¹)
		N ₂ velocity (up) (ms ⁻¹)
		NO velocity (up) (ms ⁻¹)

Table 8. NRLMSISE-00, TIE-GCM and GITM model (solar proxy) inputs.

NRLMSISE-00	TIE-GCM	GITM
81 day average F10.7	81 day average F10.7	
Daily F10.7	Daily F10.7	Daily F10.7
Ap	Ap	
		Hemispheric power index

6.3.1. CHAMP

The performance of each model is compared against the density fields derived from the CHALLENGING Minisatellite Payload (CHAMP) satellite [Reigber *et al.*, 2002, 2003]. CHAMP was in operation from July 2000 to September 2010 and the reported neutral densities are derived from accelerometer data as described by Sutton [2009]. At launch CHAMP had an orbital period of roughly 90 minutes and recorded neutral densities approximately every 45 seconds. The time series of the derived neutral densities for this test are shown in Figure 28.

For the test the observational data from CHAMP was restricted to the closest matching time to the model time series. The 30 minute time step is approximately one third of the CHAMP orbital period, which is the cause of the apparent oscillation in the CHAMP neutral densities (Figure 28).

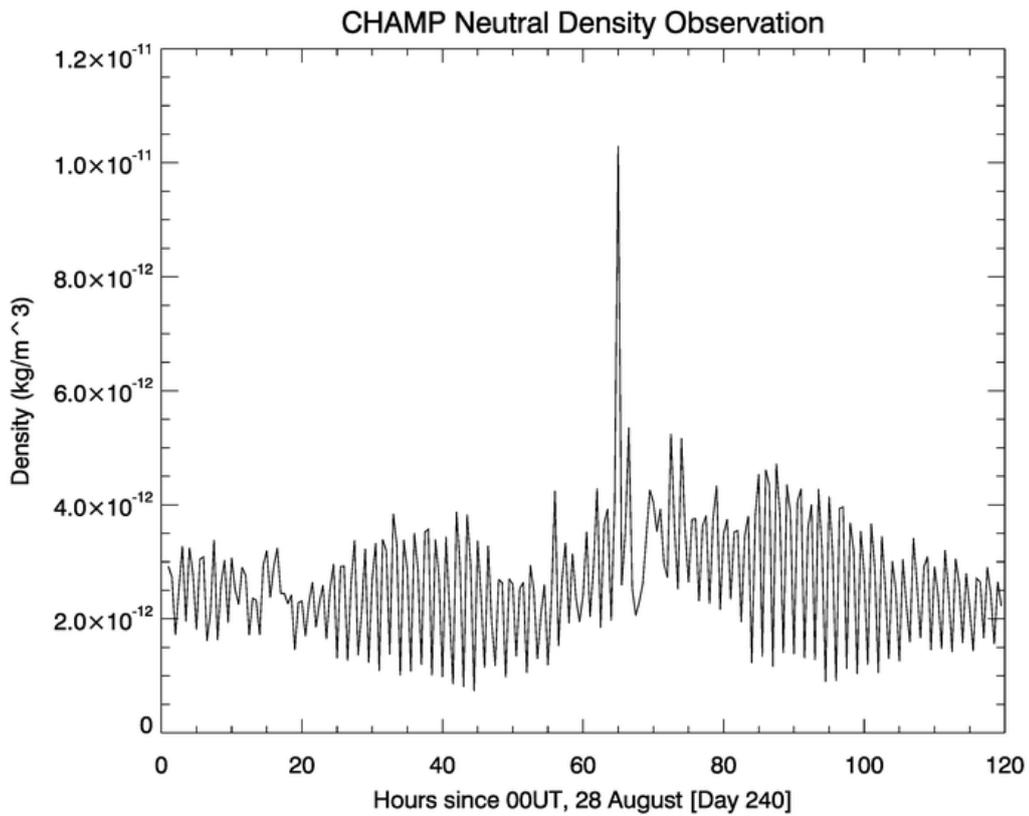


Figure 28. Neutral densities, derived from CHAMP accelerometer data for August 28th – September 1st 2009. The time step of the x-axis is 30 minutes, which was approximately one third of the CHAMP orbital period. This is the cause of the apparent oscillation in the neutral densities.

6.4. Model Comparison

To compare NRLMSISE-00, TIE-GCM and GITM with CHAMP, the output of each model was mapped to the CHAMP position using trilinear interpolation. Figure 29 shows the modified Taylor diagram for total neutral density for NRLMSISE-00, GITM and TIE-GCM compared to the CHAMP observations. Figure 30 is the time series plot of neutral density of the models and CHAMP.

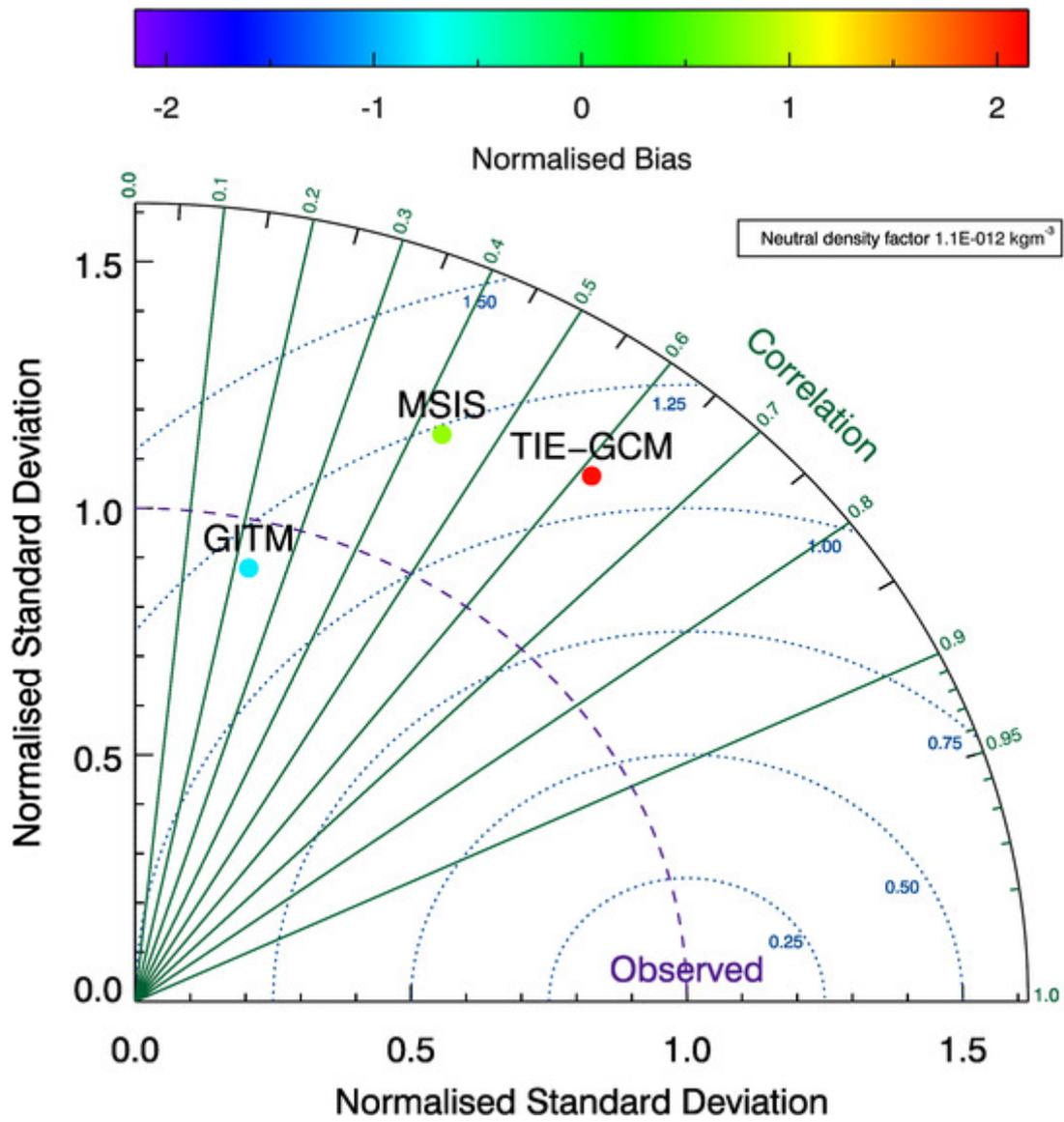


Figure 29. Modified Taylor diagram for NRLMSISE-00 (MSIS), TIE-GCM and GITM for neutral density, compared with CHAMP. The details of how to read the diagram are described in Figure 14.

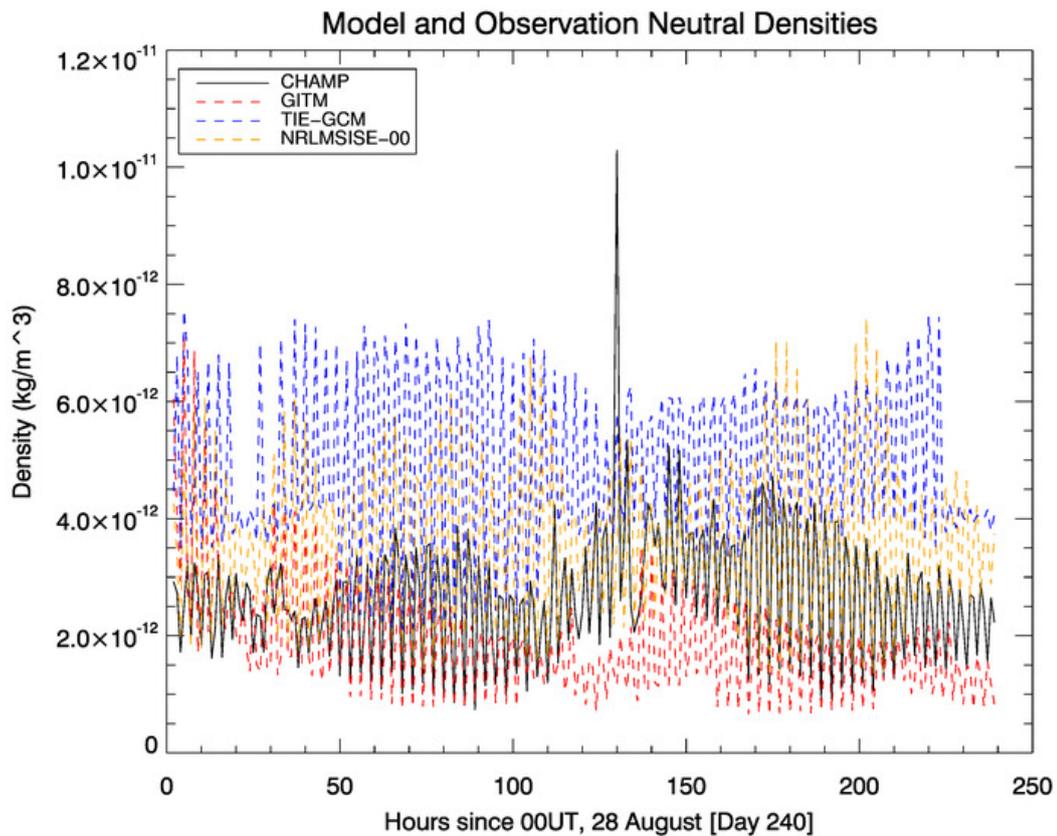


Figure 30. Plot showing the neutral density observation data from CHAMP (black) with the three model outputs, NRLMSISE-00 (orange), GITM (red) and TIE-GCM (blue) for the study period.

The NRLMSISE-00 empirical model results, as expected, show a good mean approximation to the observed state. However the model shows a larger variability in its output than the CHAMP observations. GITM shows a slight negative bias but has a standard deviation value close to the observations, i.e. the range of values that GITM produces have a similar range to the observations. Overall GITM performs the best of the three models in terms of the model mean, standard deviation and RMS values but worse in terms of correlation. GITM seems to show some reaction to the storm, with a noticeable increase in neutral density just after the peak in observed neutral densities (Figure 30). TIE-GCM has a more pronounced discrepancy between model results and observation. It has a positive bias and a standard deviation much larger than

that of the observations. However it does have the strongest correlation of the models and shows some reaction to the storm. Although there is no increase in the maximum reported values, there is an increase in the minimum values (Figure 30).

The results show that the models suffer from errors and biases, and are unable to exactly match the observed density field from CHAMP. In particular, all three models underestimate the peak during the storm. Therefore, in order to provide better forecasting abilities, techniques can be used to combine the model output to minimize the impact of model errors and bias.

6.5. Rationale of MMEs

The idea of improving model forecasts by combining two or more independent models is based upon a short note by *Thompson* [1977]. Since then, MMEs have been extensively used in the climatology community with great success [*Tracton and Kalnay*, 1993; *Harrison et al.*, 1995; *Vislocky and Fritsch*, 1995; *Doblas-Reyes et al.*, 2000, 2005; *Evans et al.*, 2000; *Fritsch et al.*, 2000; *Palmer and Shukla*, 2000; *Palmer et al.*, 2000; *Hagedorn et al.*, 2005].

An MME works on the idea that model forecasting can be improved by combining independent models [*Thompson*, 1977]. Model errors arise in a variety of forms and include computational errors in physics model solvers [*Hagedorn et al.*, 2005]. For example, many well understood physical systems use series of partial differential equations to describe that system. Yet, in order to solve them, they have to be reduced to finite-dimensional ordinary differential equations to be integrable on a computer. Whilst necessary, this reduction introduces inaccuracies.

In order to understand why an MME approach is useful, it is helpful to consider three example scenarios. Figure 31, based on the cartoon in *Hagedorn et al.* [2005], is a simplified visualization of the scenarios for two models (the red and black lines). Scenario 1 is where the models lie above and below the true state at time T and an equally weighted MME would provide the best result. In scenario 2 one of the models (black) provides the best estimate, and the MME can only be worse. However, the MME would still be better than the other model (red). Also, in practice, models provide multiple output parameters, and different models can be better at different aspects. Finally, in the 3rd scenario, neither model is very good, and again the MME would be better than one model, but worse than the other.

It is important to note that it is impossible for the MME to be worse than all of the individual models [*Hagedorn et al.*, 2005].

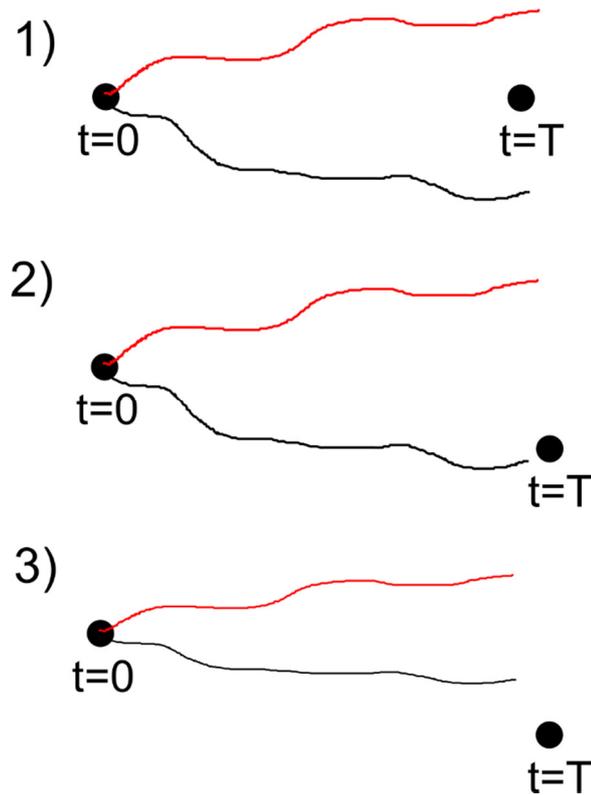


Figure 31. Idealized visualization of the basic MME scenarios, based on the cartoon in Hagedorn et al. [2005]. Two models are shown (red and black lines) for when 1) the MME would produce the best estimate, 2) one model provides the best estimate and 3) no model provides a good estimate.

However, it is clear that the MME cannot give a result better than the best model in all circumstances. For a hypothetical perfect model of a system the MME would always add worse information. However in reality such perfect models do not exist and successful MMEs should look to use independent, skilful models. It is important to use independent models since models with similar error characteristics can find such characteristics amplified in the MME. Also, if one model is shown to consistently perform worse than all other models, then this should be excluded from the MME as it does not add useful information [Hagedorn et al., 2005].

The main issue with equally weighted ensembles is that in all cases other than the first example scenario in Figure 31 they perform worse than the one best model. To overcome this problem, weighted MMEs can be constructed [*Krishnamurti, 1999; Pavan and Doblas-Reyes, 2000; Rajagopalan et al., 2002*]. However constructing model weights can be difficult, especially with small sample sizes [*Kharin and Zwiers, 2002; Hagedorn et al., 2005*].

6.6. Constructing an MME

6.6.1. Equally Weighted MME

There are a number of difficulties in constructing an MME [*Krishnamurti, 1999; Pavan and Doblas-Reyes, 2000; Rajagopalan et al., 2002*], including the fact that different models do not all share common output variables. Another problem is that there may not be observational data for each parameter, making it difficult to assess model performance for all parameters. One way to resolve the latter problem is to not take model performance into account and use an equally weighted average. Even such a simple method has been shown to increase model skill [*Barnston et al., 2003; Palmer et al., 2004; Hagedorn et al., 2005; Weisheimer et al., 2009*].

6.6.2. Weighted MME

Alternatively, the MME can use different weights for each model. There are different approaches for estimating the weights to be applied to individual models [*Krishnamurti et al., 2000; Pavan and Doblas-Reyes, 2000; Rajagopalan et al., 2002; Tebaldi and Knutti, 2007*]; however all depend on some measure of model skill. For example *Tebaldi and Knutti [2007]*

state that the skill of a (climate) model should not be judged from its ability to predict the future, but instead from its ability to predict mean conditions, variability, and transient changes.

In the absence of existing MME work in the ionospheric literature, the mean square error has been used,

$$Skill = MSE = (\mu^2 + \sigma^2). \quad (6.1)$$

Where μ is the mean of the time series of errors (model time series minus observation time series) and σ the standard deviation of the same.

6.7. Using the MME for Improved Atmospheric Density Modelling

6.7.1. MME Construction

The weighting method described in Section 6.6.2 was developed based on the models' mean conditions and variability. The mean square error (MSE; Equation (6.1)) of the models' neutral density time series compared to the CHAMP observations were used to judge model skill. Before calculating the MSE the model time series were restricted to times of low geomagnetic activity. *Fuller-Rowell and Rees* [1981] define quiet geomagnetic conditions as when the Kp index is between 0 and 1. In this study an Ap value of 3 was chosen, which corresponds to a Kp of 1-. However, restricting the time series greatly reduces the number of data points (from 240 to 50). This means the weights may not be generally applicable to the full time series [*Hagedorn et al.*, 2005].

Figure 32 is the modified Taylor diagram of NRLMSISE-00, TIE-GCM and GITM when the time series is restricted to low geomagnetic times. It is clear that each of the models perform better than for the whole time series (Figure 29). This is especially clear for the correlation between the model and observation time series. Each of the models still show their respective biases and the difference in standard deviation of each model is greater than previously.

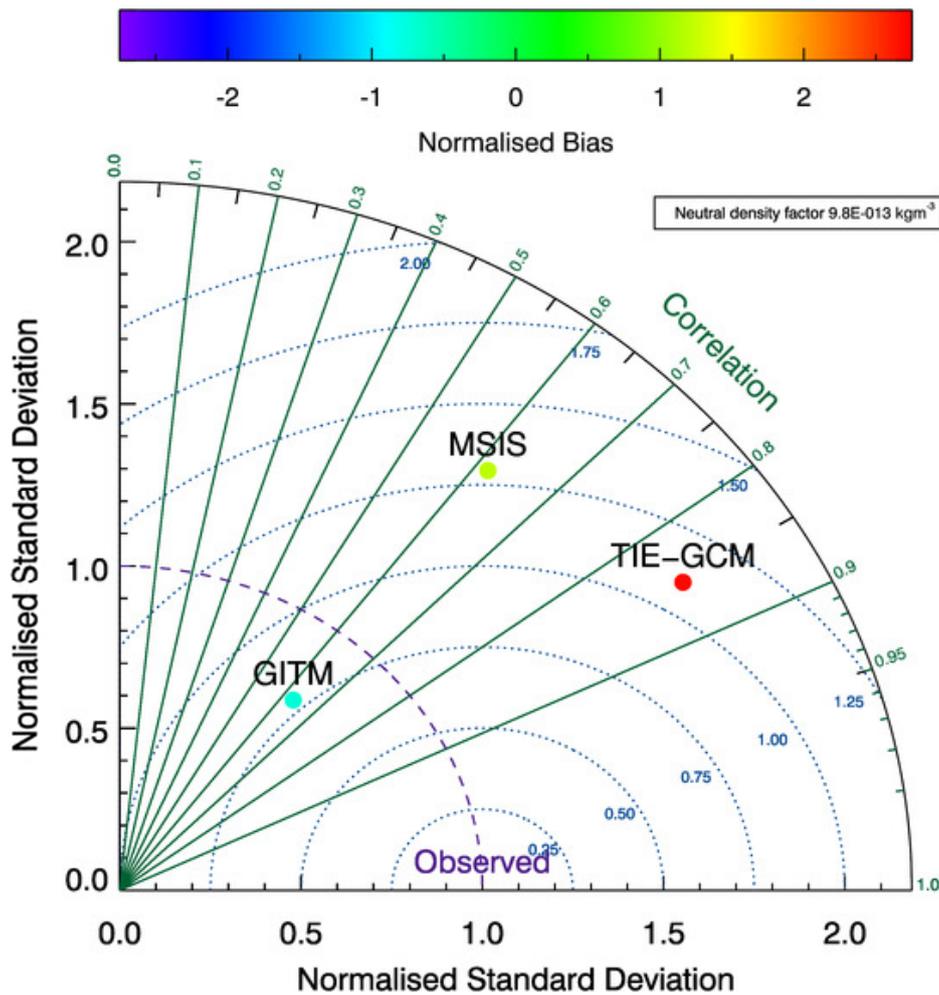


Figure 32. Modified Taylor diagram of NRLMSISE-00 (MSIS), TIE-GCM and GITM for the neutral density test. The model time series have been restricted to times when the A_p was less than or equal to 3. These time series are used to generate the weighted MME.

The model weights were based upon the model skill (Equation (6.1)) at specifying total neutral density at quiet times. The inverse of the model skill was used to weight the models, so that the model with the lowest MSE was weighted most heavily. That is, given the model skill of NRLMSISE-00, GITM and TIEGCM, S_M , S_G and S_T respectively, the weighting of model i was calculated using:

$$\text{Weighting of model } i = \frac{1}{S_i \left(\frac{1}{S_M} + \frac{1}{S_G} + \frac{1}{S_T} \right)}. \quad (6.2)$$

The model skill and weighting are given in Table 9.

Table 9. Model skill and associated weighting (calculated by the inverse of model skill, Equation (6.2)) for use in the weighted MME.

	Model Skill	Weight
NRLMSISE-00	3.03×10^{-24}	23.7%
GITM	1.06×10^{-24}	67.5%
TIE-GCM	8.15×10^{-23}	8.80%

Figure 33 and Figure 34 show the neutral density time series of the observations, average and weighted MME. Figure 35 is the modified Taylor diagram for NRLMSISE-00, GITM, TIE-GCM and the average and weighted MME. Using the figures it is clear that the average MME performs better than all of the individual models. The weighted MMEs statistics are closer to the observations for everything but correlation, in which TIE-GCM is the best. However the

difference in correlation between the equally weighted MME and TIE-GCM is not statistically significant (using Fisher's r-to-z transformation discussed in Section 5.4.1, with a Z-score of 1.161). This is due to the relatively small sample size; a five day study, with a time step of 30 minutes, results in just 240 data points (compared to 2880 used in the study in Chapter 5).

The average MME however does show some bias, whereas the weighted MME does not. The weighted MME provides variability very close to that of the CHAMP observations. However the storm period is not modelled as well as the equally weighted average as it tends to underestimate the density. This is probably due to the fact that GITM tends to underestimate the true state during the storm, whilst being heavily favoured in the weighting scheme (67.5%), which is based on quiet times (Figure 32).

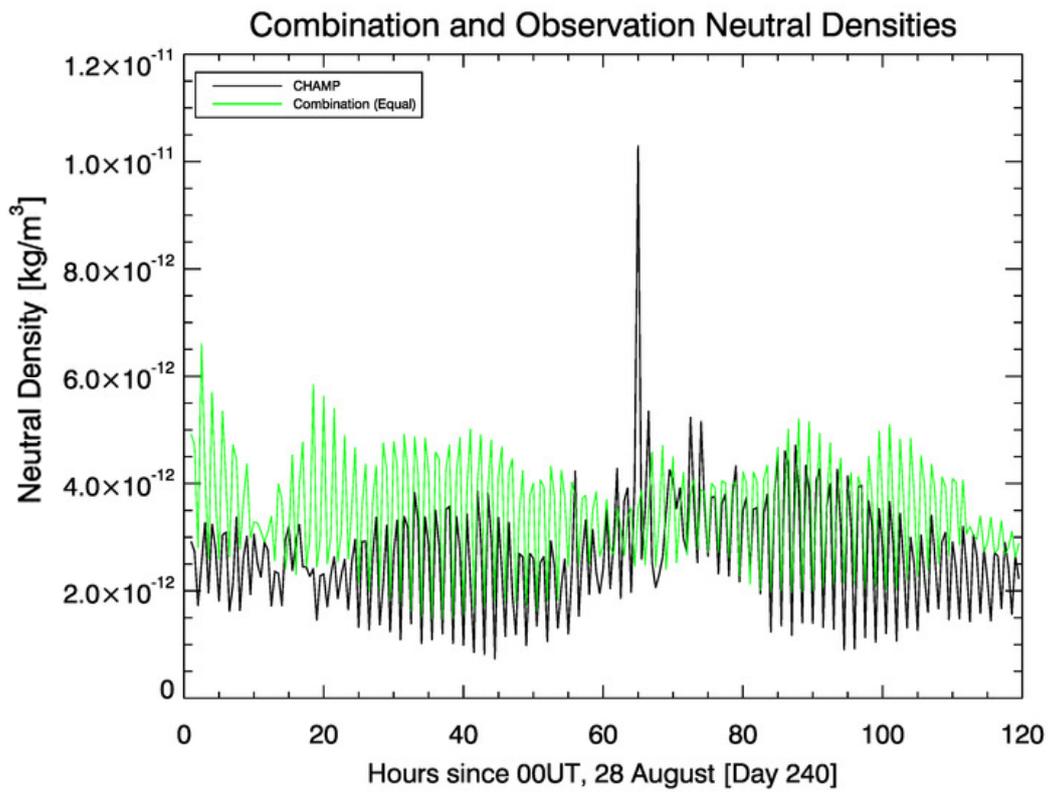


Figure 33. The CHAMP derived neutral densities are shown in black. In green are the neutral density values found from the average MME.

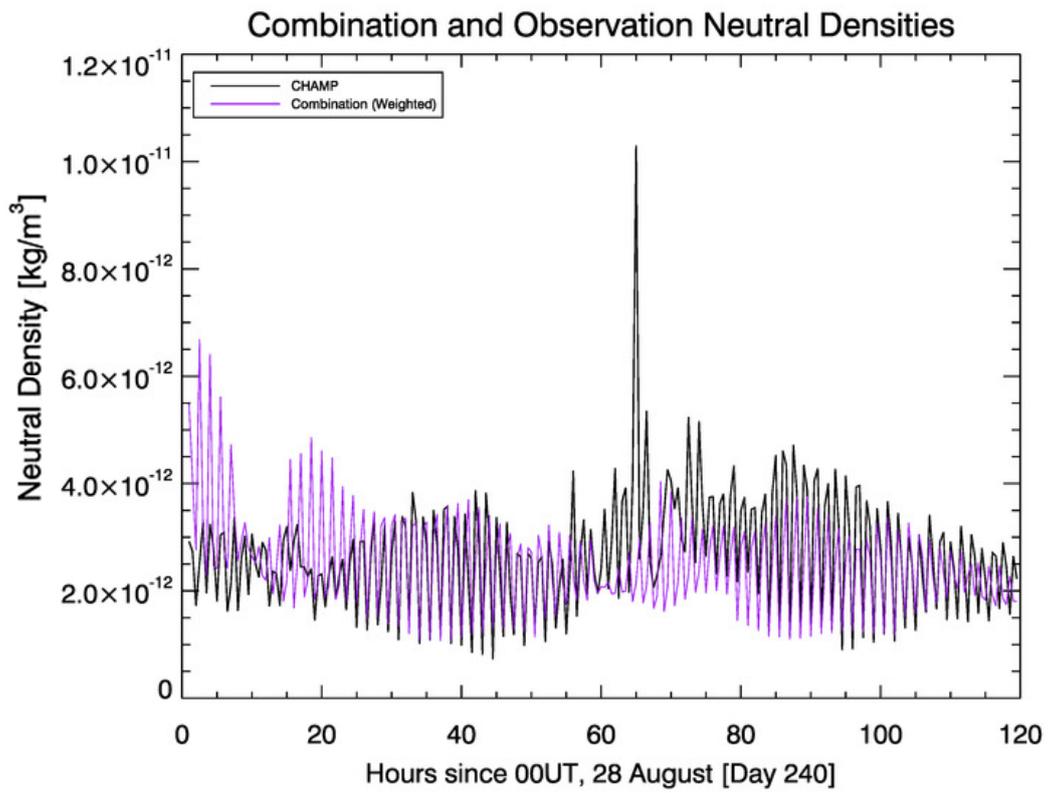


Figure 34. The CHAMP derived neutral densities are shown in black. In purple are the neutral density values found from the weighted MME.

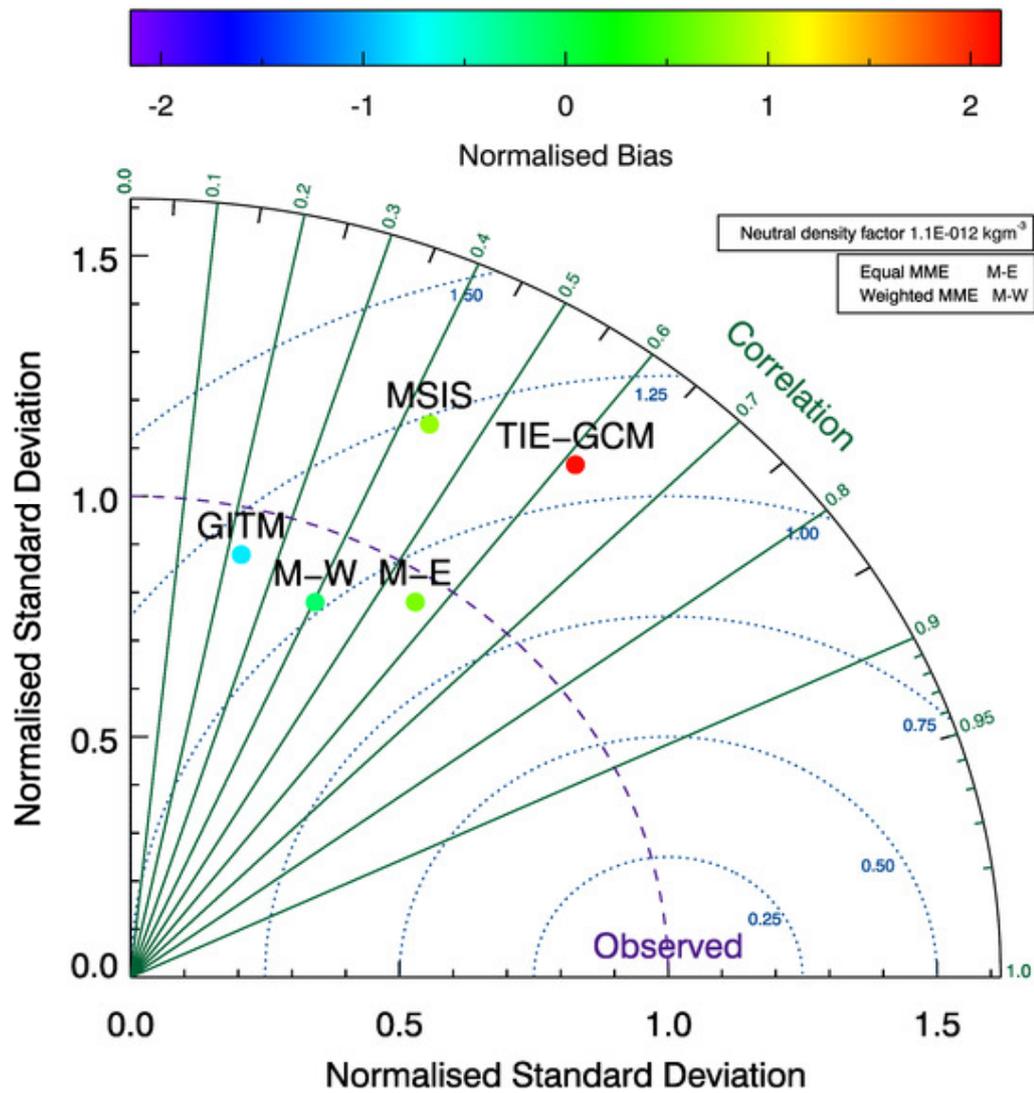


Figure 35. Modified Taylor diagram for NRLMSISE-00 (MSIS), TIE-GCM, GITM and TIE-GCM started with an MME for neutral density, compared with CHAMP. The details of how to read the diagram are described in Figure 14.

It has been shown, in this example, that combining model results leads to increased skill at matching the CHAMP derived data. In the following section this reduced uncertainty in atmospheric densities is used to provide the initial conditions of a forecast run of a model. This

approach has been previously shown to increase model forecast skill [Tebaldi and Knutti, 2007].

6.7.2. Using the MME as the Initial Conditions for TIE-GCM

The main objective of this section is to reduce the uncertainty in the initial conditions and thus increase the forecast skill of TIE-GCM. In order to use an MME as the initial conditions for a physics-based model (i.e. TIE-GCM) more than just the combined neutral density is required. The MME of each density required by TIE-GCM (Table 7), has to be calculated. Where possible the density for each model species required by TIE-GCM (i.e. oxygen; *O*) was found by combining the densities from NRLSMSE-00, GITM and TIE-GCM. However, for certain species (i.e. nitrogen oxide; *NO*) not all the models provide a density (in this case NRLMSISE-00). In these cases, just the models which do provide a density value were used. In case TIE-GCM has a density which no other model provides, the original data is used on its own. A similar approach is used for the temperatures and velocities.

To combine densities, temperatures and velocities from multiple models the data must be interpolated to common latitude, longitude and altitude grids. To achieve that, the NRLMSISE-00 and GITM grids were trilinearly interpolated to the TIE-GCM grid. The grids were then combined to form an MME.

For the new TIE-GCM run the model was restarted using the MME grid as the initial conditions. TIE-GCM was then run for six hours with the model output recorded every 30 minutes. After this period TIE-GCM was again restarted using the MME grid for the next six hour period. For this forecast run the model only used Ap and F10.7 in the initial conditions. It was not updated

at each time step. This was so a true forecast could be simulated. Figure 36 is a flow chart of the process used to run TIE-GCM with the MME as its initial conditions for a six hour forecast, and Figure 37 is the procedure used for this test scenario.

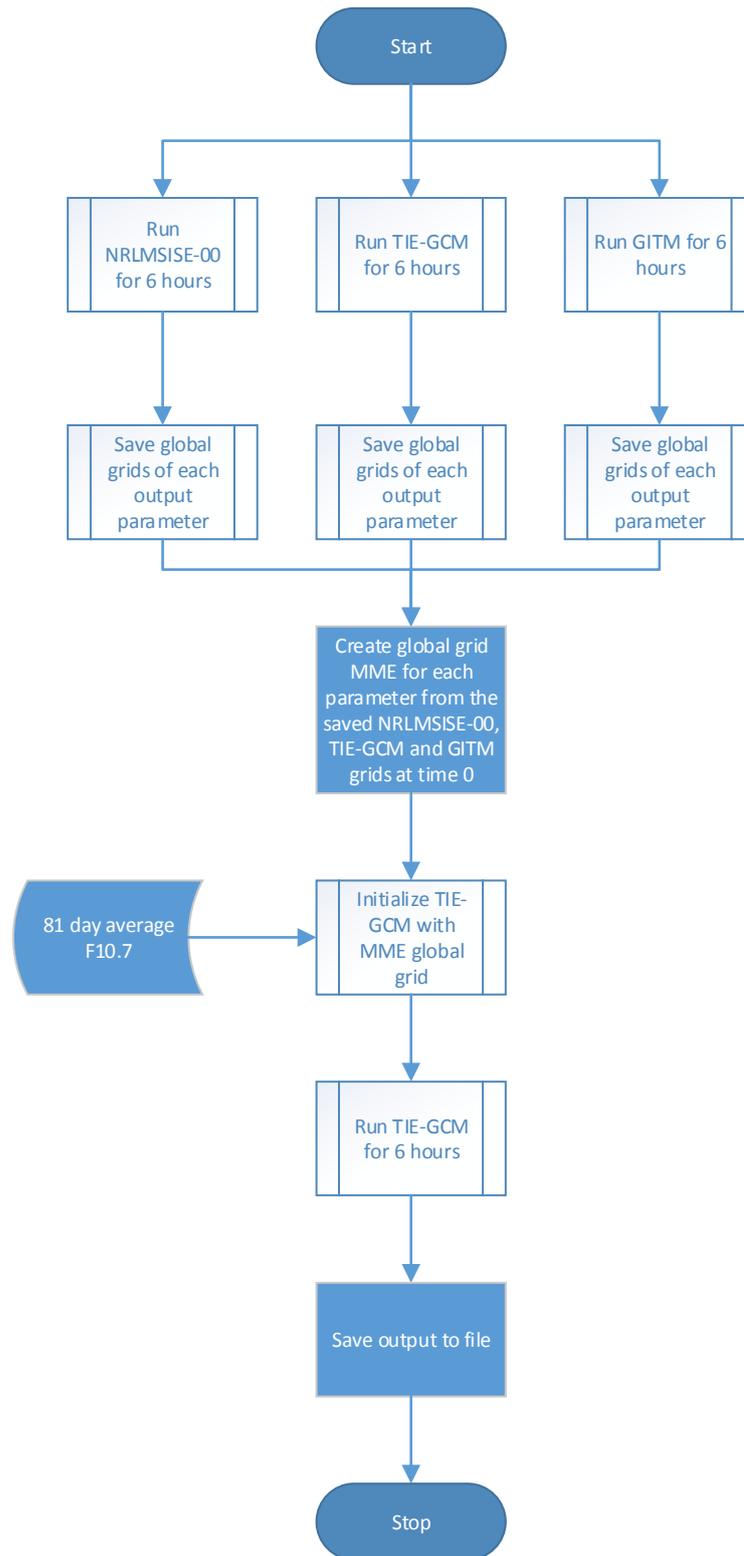


Figure 36. Flow chart of the procedure for running TIE-GCM using the MME as its initial conditions for a six hour forecast.

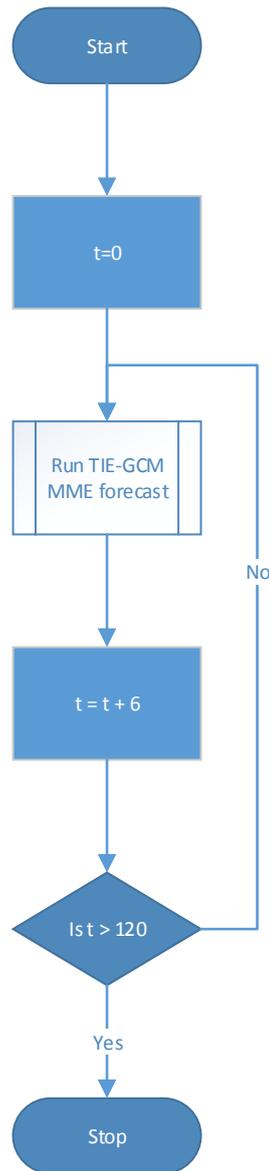


Figure 37. Procedure for finding the TIE-GCM forecast using an MME as its initial conditions for the August 28th – September 1st 2009 neutral density study. The “run TIE-GCM MME forecast” process refers to the procedure described in Figure 36.

Figure 38 shows the modified Taylor diagram for neutral density compared to the CHAMP observations for the original TIE-GCM run, the NRLMSISE-00 (MSIS), GITM results and the results of rerunning TIE-GCM using both the average and weighted MME as the initial

condition every six hours. Figure 39 is the reported time series of neutral densities from the CHAMP observations, the original TIE-GCM run, the equally weighted combination MME and the results of rerunning TIE-GCM using the MME.

Using the MME as the initial condition in TIE-GCM provides a clear improvement in all statistical parameters, compared to any of the contributing models. Initially the model is well above the observations (Figure 39), but this is due to the large initial values in the MME (caused by the GITM ‘spin up’ time, along with the already large TIE-GCM and NRLMSISE-00 values). However after approximately 20 hours the reported densities show very low bias and have variability close to the observations. In particular, the post-storm period is modelled very accurately. The average MME initial conditions for TIE-GCM run does not significantly improve upon the original TIE-GCM correlation, but has significantly improved the bias and standard deviation of the model. The new TIE-GCM run (using the average MME) offers an improvement (in bias and correlation) upon the neutral density MME calculated after the models were run (Figure 35). This is since the physics of one model, given initial conditions with lower errors, can propagate densities better than the average of three models, each with poor initial conditions.

None of the contributing models, nor the MMEs, model the peak of the storm period (~65 hours after August 28th 2009) with any accuracy. The best the models can do is to try and model the post-storm period as best as possible. This is since the models do not react as sharply to the increase in A_p (in terms of reported neutral densities) as compared to reality.

Using the MME as the initial conditions to TIE-GCM offers a 55% reduction of the RMS error compared to the initial TIE-GCM runs from Section 6.4.

Figure 38 also includes the results from using the weighted MME as the initial condition every six hours in TIE-GCM. The time series plot is shown in Figure 40. These figures show that using the weighted combination MME as the initial condition for TIE-GCM also provides significant improvement. Before the storm onset the new TIE-GCM run gives a non-biased result with a standard deviation close to that of the observations. However directly after the storm the output does not follow the density values so well; for the remaining 50 hours (from hour 70 to 120) the TIE-GCM result shows a smaller variability than that of the observations.

Compared to the TIE-GCM initialised with the average MME the weighted MME performs much worse. Although the bias, and variability, is better than any of the contributing models the correlation is poor. Indeed, the correlation is very similar to the most heavily weighted model, GITM. However using the weighted MME as the initial condition still offers considerable improvement, especially with regards to the parameter that has been used to judge the model skill, the mean square error. The weighted MME TIE-GCM run gives a 48% reduction in the RMS error compared to the original TIE-GCM run.

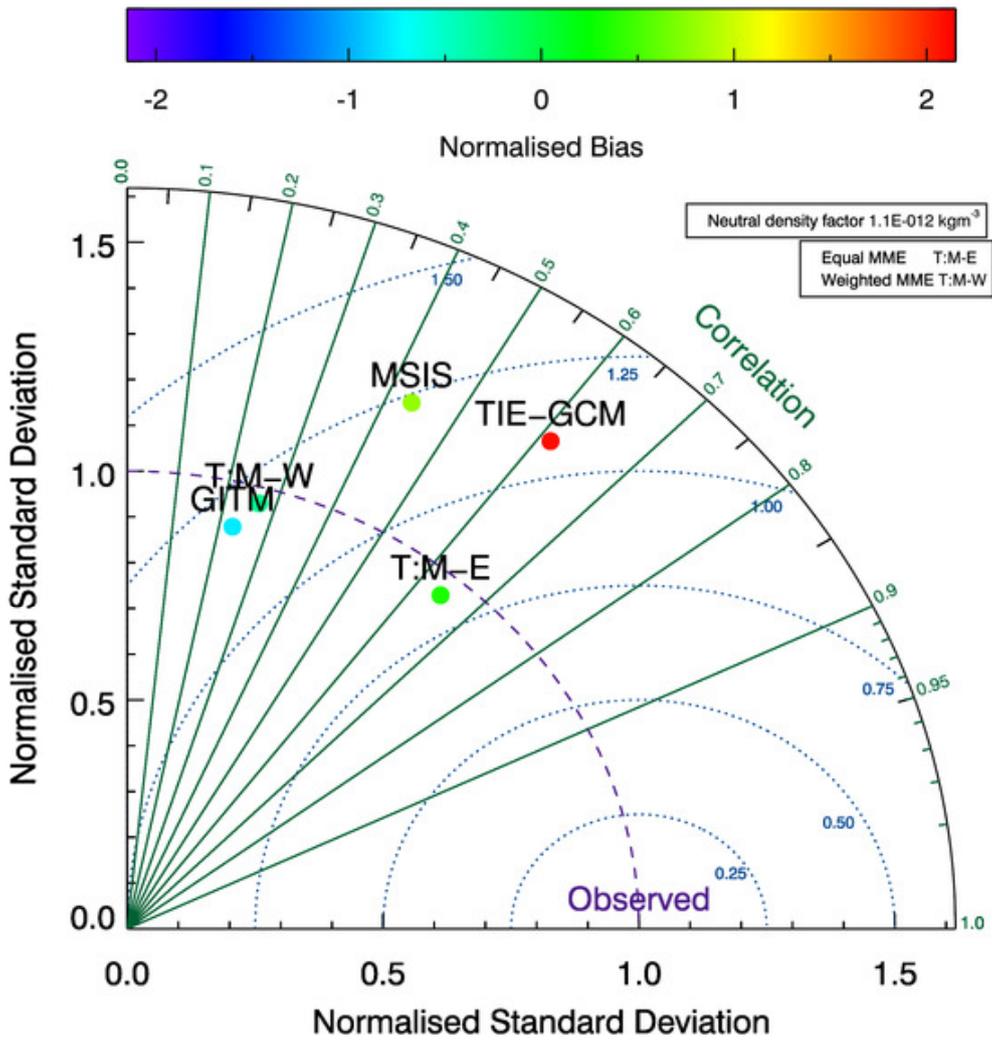


Figure 38. Modified Taylor diagram for NRLMSISE-00 (MSIS), TIE-GCM, GITM and for TIE-GCM using the MME (both average and weighted) for its initial conditions every six hours, compared with the CHAMP observations. The details of how to read the diagram are described in Figure 14.

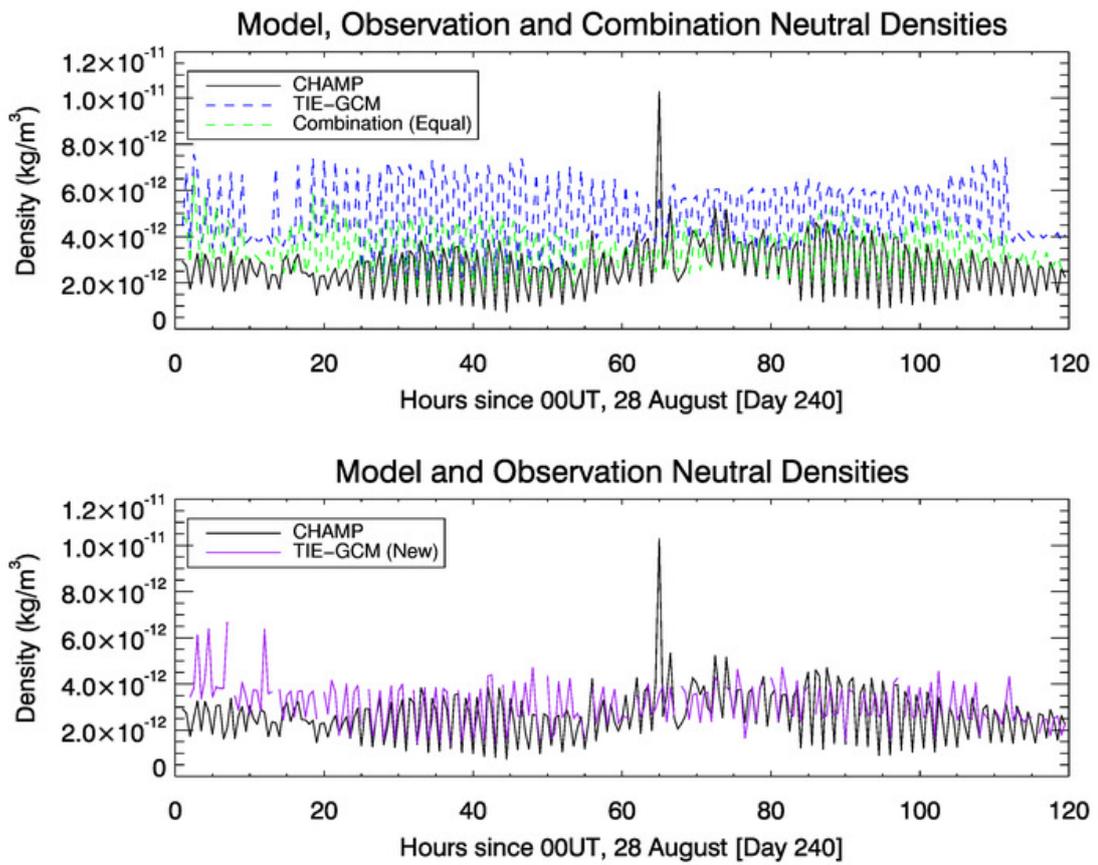


Figure 39. Top panel shows the neutral density from the CHAMP observations, from the original TIE-GCM run and the equal combination MME. The bottom panel shows the CHAMP observations and the new TIE-GCM output, using the MME as the initial condition every six hours.

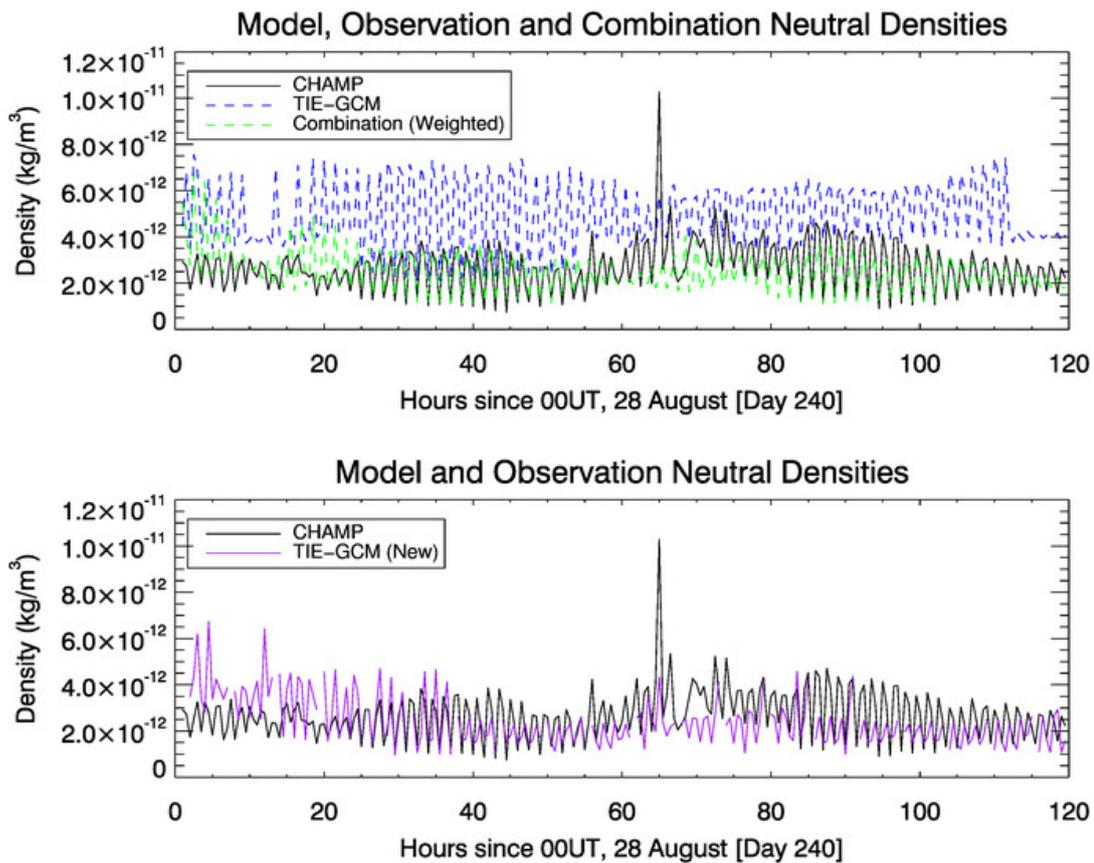


Figure 40. Top panel shows the neutral density from the CHAMP observations, from the original TIE-GCM run and the weighted MME. The bottom panel shows the CHAMP observations and the new TIE-GCM output, using the MME as the initial condition every six hours.

6.8. Conclusions

The work presented in this chapter shows the possibility of using multi-model ensembles (MMEs) to enhance the forecast skill of ionosphere-thermosphere models. Three models were used: an empirical model (NRLMSISE-00) and two physics-based models (TIE-GCM and GITM). The models' output density has been compared against density fields from CHAMP,

where all models have large errors compared to the observations. To improve the density estimation, an MME averaging technique has been applied and tested. Two approaches for the MME were used, a simple average MME where all models have the same weight, and a weighted MME, where each model is weighted according to its skill. The results show a significant improvement in both cases. The MME was then used to initialize one of the physics-based models (TIE-GCM) to try and improve its forecast skill. As with the previous case, the first experiment uses an MME with a simple average as the initial conditions to TIE-GCM. This initialization shows a 55% reduction in RMS error over the TIE-GCM run. The weighted MME performs worse than the simple ensemble average MME, offering an approximate 48% reduction in RMS error. There are a number of reasons for this worse performance, including that the weighting scheme uses the model quiet time to generate the model weights which is then applied across all time periods. This is consistent with *Hagedorn et al.* [2005] who argued that for small data sets the only way to generate a good MME is to use the average.

A possible improvement would be to use two different weighting schemes, one for quiet times and one for storm times. Alternatively a weighting scheme that is dependent on location as well as conditions could be beneficial. A further approach would be to change the weighting scheme altogether and adopt Reliability Ensemble Averaging (REA) that is often used to generate MMEs in climatology studies [*Giorgi and Mearns, 2002*]. Also, in order to verify the results that have been presented here longer test scenarios should be used to reduce the uncertainties in the statistics.

7. CONCLUSIONS AND FUTURE WORK

*I do not mind if you think slowly, but I do object
when you publish more quickly than you can think.*

- Wolfgang Pauli

Space weather can affect RF systems in many ways. Affected systems include global navigation satellite systems (GNSS), high frequency (HF) communications, space-based Earth observation radars and space situational awareness radars. Ionospheric models are often used to mitigate negative effects.

In order to characterise individual model performance, and thereby choose the best model for mitigating space weather effects, suitable metric(s) and common testing scenarios are required. A review of current model comparison techniques both in the ionospheric literature, and wider scientific community has been discussed. Problems with current methods of model comparison and with the visualization of model results, in the ionospheric literature, have been described. A method to overcome a number of these problems as been suggested, in the form of modified Taylor diagrams.

Taylor diagrams have been used in the climatology and weather community since their introduction in 2001. However no reference to their use in the ionospheric literature could be found. Taylor diagrams do however have two main problems which the modifications presented here have addressed. First, Taylor diagrams do not show if a model has any bias, which is an important consideration when comparing model performance. Secondly, Taylor diagrams use the little known “centred-pattern RMS difference”, the mean-removed RMS difference. However it is not well understood and in a number of publications it is incorrectly used

interchangeably with the RMS difference. It has, however, been shown that the centred-pattern RMS difference is equivalent to the standard deviation of the model errors (model time series minus observation time series). This is a more common modelling parameter, and avoids some confusion. The modified diagrams also allow one to find the full mean square error by adding in quadrature the model bias and the error standard deviation. Finally, the normalisation factor has been included in the diagrams to allow the reader to reconstitute statistical values which have been normalised in order to plot multiple parameters on one diagram.

Using modified Taylor diagrams IRI, EDAM, GPSII, the data ingestion version of NeQuick (mNeQuick), NRLMSISE-00, TIE-GCM and GITM have been compared. This is the first known comparison of the three main model types (empirical, data assimilation and physics-based) for a long test study (one month). The comparison discussed the importance of stationary time series before calculating statistics. Using modified Taylor diagrams the data assimilative models, EDAM, GPSII and mNeQuick have been shown to offer significant improvements in foF2 and h(0.8foF2) when compared to IRI. For hmF2, GPSII and mNeQuick improve upon IRI. EDAM, however, does not statistically improve upon IRI. It has been shown, though, that EDAM is not defaulting to its background model (IRI-2007) and is making changes to its electron density grid. TIE-GCM has been shown to perform comparably to IRI in terms of hmF2, albeit with a large positive bias, but worse in both foF2 and h(0.8foF2). However, the model performs sufficiently well that it is believed that the model could be used as background model for future data assimilation models. GITM performed the worst of the tested models, which is at least partly due to the unphysical electron densities which the model produces at night.

The ideas of combining the model parameter time series and electron density profiles were also discussed. It has been shown that these two simple combination methods provide significant improvements over the models used to construct the combination. It was this improvement which inspired using model combinations, or ensembles, as initial conditions in forecast runs of the physics-based model, TIE-GCM.

The suitability of multi-model ensembles (MMEs) for use in ionospheric/thermospheric modelling has been described. The two main approaches for constructing them have been demonstrated. For a neutral density study during solar minimum, MMEs have been shown to offer statistical improvements over the contributing models, when their outputs are compared to CHAMP observations. These improved conditions have then been used as the initial conditions for TIE-GCM. It has been shown that using the average (i.e. equally weighted) MME caused a 55% reduction in the RMS error for neutral densities. The weighted MME exhibited less of an improvement. The RMS error was reduced by 48%, but the model results showed very poor correlation with the observation.

Future work should be carried out on a number of fronts. First, more standardized testing scenarios should be published, allowing models to be further tested and compared. This will help to highlight which modelling approaches are most useful for operational RF systems. Secondly, for MMEs a variety of work should be done to advance the field. Longer test studies (on the order of a month), would increase the confidence of the statistics. Increasing the number of models used to construct the MME would also help to reduce the errors in densities even further, potentially leading to better forecasts. Finally, addressing the problems of the weighted MME should be undertaken. The weighting scheme could be based on more times, rather than

just quiet times. Also, it could be varied spatially. Furthermore, other weighting techniques such as reliability ensemble averaging (REA) (currently used in the climatology literature) could be used.

APPENDIX A. THE RELATIONSHIP BETWEEN ELECTRON DENSITY AND PLASMA FREQUENCY

The Appleton-Hartree-Lassen equation describes the refractive index for electromagnetic wave propagation in a cold, magnetized, plasma [*Lassen, 1927; Hartree, 1929; Appleton, 1932; Davies, 1990*]. The refractive index (n) can be calculated using the simplified form (no collisions) of the Appleton-Hartree-Lassen equation [*Angling et al., 2013*]:

$$n^2 = 1 - \frac{2X(1-X)}{2(1-X) - Y_T^2 \pm \sqrt{Y_T^4 + 4(1-X)^2 Y_L^2}} \quad (\text{A.1})$$

Where

$$X = \frac{N_e q_e^2}{\epsilon_0 m \omega^2}, \quad (\text{A.2})$$

$$Y_T = \frac{q_e B_T}{m \omega}, \quad (\text{A.3})$$

$$Y_L = \frac{q_e B_L}{m \omega}, \quad (\text{A.4})$$

and N_e is the electron density, q_e is the charge of the electron (1.6×10^{-19} coulomb), ϵ_0 is the vacuum permittivity (8.85×10^{-12} farad-meter⁻¹), m is the electron mass (9.11×10^{-31} kg), ω is the angular frequency, B_L is the longitudinal component of the magnetic field and B_T is the transverse component of the magnetic field. The positive part of the plus/minus sign in the denominator refers to the O wave whilst the negative part refers to the X wave.

Under the assumption of no magnetic field, $Y_T = Y_L = 0$:

$$n^2 = 1 - X, \quad (\text{A.5})$$

$$= 1 - \frac{N_e q_e^2}{\epsilon_0 m \omega^2}. \quad (\text{A.6})$$

Then, since plasma frequency is defined as:

$$\omega_p^2 = \frac{N_e q_e^2}{\epsilon_0 m}, \quad (\text{A.7})$$

Equation (A.7) can be substituted into Equation (A.6) to give:

$$1 - \frac{N_e q_e^2}{\epsilon_0 m \omega^2} = 1 - \frac{\omega_p^2}{\omega^2}. \quad (\text{A.8})$$

Converting angular frequency to frequency ($\omega = 2\pi f$) and rearranging the formula, the conversion between plasma frequency and electron density can be arrived at:

$$N_e \cdot \frac{q_e^2}{4\pi^2 \epsilon_0 m} = f_p^2. \quad (\text{A.9})$$

That is, the electron density (in electrons/m³) multiplied by a constant, is equal to the square of the plasma frequency (in Hz). This constant is $\sim 80.616 \text{ m}^3\text{s}^{-2}$, which is often approximated to 81 and Equation (A.9) is usually written thus:

$$f_p = 9\sqrt{N_e}. \quad (\text{A.10})$$

A more detailed discussion of the Appleton-Hartree-Lassen equation can be found in *Davies* [1990].

APPENDIX B. ESTIMATION OF TEC FROM GNSS MEASUREMENTS

In a dispersive medium, such as the ionosphere, the angular frequency, ω , and angular wave number, k , are not proportional. The dispersion relationship is defined as [Crawford, 1968]:

$$\omega^2 = c^2 k^2 + \omega_p^2, \quad (\text{B.1})$$

where c is the speed of light and ω_p the angular plasma frequency. From Section 2.6, $\omega_p = 2\pi f_p$ and $f_p^2 = 80.6N_e$. These results can be combined with the phase and group velocity to find the total electron content (TEC). The phase and group velocity are defined as [Xu, 2007]:

$$\text{Phase Velocity} = v_{ph} = \frac{\omega}{k}, \quad (\text{B.2})$$

$$\text{Group Velocity} = v_{gr} = \frac{\partial \omega}{\partial k}. \quad (\text{B.3})$$

Converting angular frequency to frequency ($\omega = 2\pi f$),

$$v_{ph}^2 = \frac{\omega^2}{k^2}, \quad (\text{B.4})$$

$$= \frac{4\pi^2 f^2}{k^2}. \quad (\text{B.5})$$

Rearranging Equation (B.1) in terms of k^2 , converting to frequency and substituting into Equation (B.5):

$$v_{ph}^2 = \frac{4\pi^2 f^2}{\left(\frac{\omega^2 - \omega_p^2}{c^2}\right)}, \quad (\text{B.6})$$

$$= \frac{4\pi^2 f^2}{\left(\frac{4\pi^2 f^2 - 4\pi^2 f_p^2}{c^2}\right)}, \quad (\text{B.7})$$

$$= \frac{c^2}{\left(1 - \left(\frac{f_p}{f}\right)^2\right)}. \quad (\text{B.8})$$

The refractive index of a medium is defined as the ratio of the speed of the wave in a vacuum to the speed in the medium (in this case the ionosphere), that is $n_{ph} = \frac{c}{v_{ph}}$ where n_{ph} is the phase refractive index. Using v_{ph} as found in Equation (B.8) n_{ph} can be written as:

$$n_{ph} = \frac{c}{\left(\frac{c}{\sqrt{1 - (f_p/f)^2}}\right)}, \quad (\text{B.9})$$

$$= \sqrt{1 - \left(\frac{f_p}{f}\right)^2}. \quad (\text{B.10})$$

Since $\frac{f_p}{f} \ll 1$, the first order Taylor expansion approximation $\sqrt{1 - \epsilon^2} \approx 1 - \frac{1}{2}\epsilon^2$ can be used.

Using $f_p = 80.6N_e$ (Section 2.6) the phase refractive index in terms of frequency and electron density can be found:

$$n_{ph} \approx 1 - \frac{1}{2} \left(\frac{f_p}{f} \right)^2, \quad (\text{B.11})$$

$$= 1 - \frac{1}{2} \left(\frac{80.6 N_e}{f^2} \right), \quad (\text{B.12})$$

$$= 1 - \frac{40.3}{f^2} N_e. \quad (\text{B.13})$$

To find the group refractive index, in terms of frequency and electron density, rearrange Equation (B.1) in terms of ω and differentiate:

$$\omega = (c^2 k^2 + \omega_p^2)^{\frac{1}{2}}, \quad (\text{B.14})$$

$$\frac{\partial \omega}{\partial k} = \frac{1}{2} (c^2 k^2 + \omega_p^2)^{-\frac{1}{2}} \cdot 2c^2 k, \quad (\text{B.15})$$

$$= \frac{c^2 k}{\sqrt{c^2 k^2 + \omega_p^2}}. \quad (\text{B.16})$$

Recall from Equation (B.3) that $v_{gr} = \frac{\partial \omega}{\partial k}$ thus:

$$v_{gr} = \frac{c^2 k}{\sqrt{c^2 k^2 + \omega_p^2}}. \quad (\text{B.17})$$

As before substitute Equation (B.1) (rearranged in terms of k^2) into Equation (B.17) to get:

$$v_{gr} = \frac{c^2 \sqrt{\frac{\omega^2 - \omega_p^2}{c^2}}}{\sqrt{\frac{c^2(\omega^2 - \omega_p^2)}{c^2} + \omega_p^2}}. \quad (\text{B.18})$$

Replacing angular frequency with frequency, and cancelling down, Equation (B.18) becomes:

$$v_{gr} = c \sqrt{1 - \left(\frac{f_p}{f}\right)^2}. \quad (\text{B.19})$$

Since the group refractive index is $n_{gr} = \frac{c}{v_{gr}}$,

$$n_{gr} = \frac{c}{c \sqrt{1 - \left(\frac{f_p}{f}\right)^2}}, \quad (\text{B.20})$$

$$= \left(1 - \left(\frac{f_p}{f}\right)^2\right)^{-\frac{1}{2}}. \quad (\text{B.21})$$

Again, since $\frac{f_p}{f} \ll 1$, the approximation $(1 - \epsilon^2)^{-\frac{1}{2}} \approx 1 + \frac{1}{2}\epsilon^2$ can be used and thus:

$$n_{gr} = 1 + \frac{1}{2} \left(\frac{f_p}{f}\right)^2, \quad (\text{B.22})$$

$$= 1 + \frac{40.3}{f^2} N_e. \quad (\text{B.23})$$

Using the refractive index one can find the difference between the measured range and the geometric range (a refractive index of 1). That is, the difference between a ray travelling along the same path but experiencing a refractive index of 1:

$$\Delta_{ph} = \int n \, dl - \int dl, \quad (\text{B.24})$$

$$\Delta_{ph} = \int (n - 1) \, dl. \quad (\text{B.25})$$

Then substituting the phase and group refractive indices from Equation (B.13) and Equation (B.23) into Equation (B.25) and using the definition of TEC from Equation (2.7):

$$\Delta_{iono,ph} = \int \left(\left[1 - \frac{40.3}{f^2} N_e \right] - 1 \right) dl, \quad (\text{B.26})$$

$$= -\frac{40.3}{f^2} \int N_e \, dl, \quad (\text{B.27})$$

$$= -\frac{40.3}{f^2} TEC. \quad (\text{B.28})$$

$$\Delta_{iono,gr} = \int \left(\left[1 + \frac{40.3}{f^2} N_e \right] - 1 \right) dl, \quad (\text{B.29})$$

$$= \frac{40.3}{f^2} \int N_e \, dl, \quad (\text{B.30})$$

$$= \frac{40.3}{f^2} TEC. \quad (\text{B.31})$$

Thus there is a phase advance when passing through the ionosphere, and a group delay.

It should be noted that $n_{gr} \geq n_{ph}$ implies $v_{gr} \leq v_{ph}$. Consequently for $n_{ph} \leq 1$, $v_{ph} = \frac{c}{n_{ph}} \geq c$, i.e. the phase velocity travels faster than the speed of light. However, no information is transmitted with the phase and as such Special Relativity is not violated.

A GNSS can measure both the phase and pseudorange for a signal passing through the ionosphere. An estimation of the TEC can be found using a linear combination of either the pseudorange or phase measurements from two frequencies, such as L1 and L2 for GPS.

Firstly, using the pseudorange, let $I_i = \frac{40.3}{f_i^2} TEC$ (i.e. $\Delta_{iono,gr}$ from Equation (B.31)) for $i = 1, 2$ and thus:

$$P_1 = \rho_1 + I_1 \quad \text{and} \quad P_2 = \rho_2 + I_2, \quad (\text{B.32})$$

where P_i is the pseudorange of signal L_i and ρ_i is the delay due to geometric range, clock bias and the troposphere for each signal. Multiplying Equation (B.32) by its corresponding signal frequency gives:

$$f_1^2 P_1 = f_1^2 \rho_1 + f_1^2 I_1, \quad (\text{B.33})$$

$$f_2^2 P_2 = f_2^2 \rho_2 + f_2^2 I_2. \quad (\text{B.34})$$

Taking the difference of Equation (B.33) and (B.34), with the assumption the L1 and L2 signals travel along the same path and as such $\rho_1 = \rho_2$, and noting that:

$$f_1^2 I_1 = f_1^2 \cdot \frac{40.3}{f_1^2} TEC, \quad (\text{B.35})$$

$$= 40.3 \cdot TEC, \quad (\text{B.36})$$

$$= \frac{f_2^2}{f_2^2} 40.3 \cdot TEC, \quad (\text{B.37})$$

$$= f_2^2 I_2, \quad (\text{B.38})$$

gives,

$$f_1^2 P_1 - f_2^2 P_2 = (f_1^2 - f_2^2) \rho. \quad (\text{B.39})$$

Rearranging to make P_2 the subject of Equation (B.39):

$$P_2 = \frac{f_1^2}{f_2^2} P_1 - \rho \left(\frac{f_1^2}{f_2^2} - 1 \right), \quad (\text{B.40})$$

Then, since $\rho = P_1 - I_1$, substitute this into Equation (B.40),

$$P_2 = \frac{f_1^2}{f_2^2} P_1 - (P_1 - I_1) \left(\frac{f_1^2}{f_2^2} - 1 \right), \quad (\text{B.41})$$

$$P_2 = P_1 + \frac{I_1 f_1^2}{f_2^2} - I_1, \quad (\text{B.42})$$

$$P_1 - P_2 = I_1 \left(1 - \frac{f_1^2}{f_2^2} \right). \quad (\text{B.43})$$

Substituting the definition of I_1 into Equation (B.43) gives

$$P_1 - P_2 = \frac{40.3}{f_1^2} TEC \left(1 - \frac{f_1^2}{f_2^2} \right), \quad (\text{B.44})$$

and thus the pseudorange TEC is given by:

$$TEC_{\text{pseudo-range}} = \frac{f_1^2 f_2^2 (P_1 - P_2)}{40.3(f_2^2 - f_1^2)}. \quad (\text{B.45})$$

Using the pseudorange TEC estimate can result in poor results since pseudorange is a noisy measurement. On the other hand using phase to calculate the TEC gives ambiguous results. The approach usually taken is to use the noisy pseudorange measurement to level the ambiguous phase TEC results.

The observed phase is defined similarly to the pseudorange (Equation (B.32)), but I_i is now defined as $-\frac{40.3}{f_i^2} TEC$ (i.e. $\Delta_{iono,ph}$ from Equation (B.28)), that is:

$$\phi_1 = \rho + I_1 + \lambda_1(2\pi n_1) \quad \text{and} \quad \phi_2 = \rho + I_2 + \lambda_2(2\pi n_2), \quad (\text{B.46})$$

Where ϕ_i is the phase and $2\pi n_i$ is the phase ambiguity. Then following the same argument as from Equation (B.33) to (B.44), phase TEC is defined as:

$$TEC_{phase} = \frac{f_1^2 f_2^2 (\phi_2 - \phi_1 - 2\pi(\lambda_1 n_1 - \lambda_2 n_2))}{40.3(f_2^2 - f_1^2)}, \quad (B.47)$$

where $2\pi(\lambda_1 N_1 - \lambda_2 N_2)$ is the bias due to the ambiguity. The bias can be rewritten as a constant and this leaves the final phase TEC equation as:

$$TEC_{phase} = \frac{f_1^2 f_2^2 (\phi_2 - \phi_1)}{40.3(f_2^2 - f_1^2)} + A, \quad (B.48)$$

where, A is the unknown ambiguity. Then, levelling the ambiguous phase TEC with the pseudorange TEC, by subtracting the mean of the TEC phase, and adding the mean of the pseudorange TEC one can find the TEC equation:

$$TEC = TEC_{phase} - \overline{TEC_{phase}} + \overline{TEC_{pseudorange}}. \quad (B.49)$$

In practice, the pseudorange TEC measurement is not without biases, usually due to receiver noise. These biases can be calculated for using a differential code bias (DCB).

LIST OF REFERENCES

Nunc est bibendum

- Horace, Odes

AchutaRao, K., and K. R. Sperber (2006), ENSO simulation in coupled ocean-atmosphere models: are the current models better?, *Clim. Dyn.*, 27, 1–15, doi:10.1007/s00382-006-0119-7.

Akaike, H. (1973), Information theory and an extension of the maximum likelihood principle, in *Second International Symposium on Information Theory*, edited by B. N. Petrox and F. Caski, pp. 267–281.

Anderson, D. N., J.M.Forbes, and M. Codrescu (1989), A Fully Analytical, Low- and Middle-Latitude Ionospheric Model, *J. Geophys. Res.*, 94, 1520.

Angling, M. J. (2008), First assimilations of COSMIC radio occultation data into the Electron Density Assimilative Model (EDAM), *Ann. Geophys.*, 26(2), 353–359.

Angling, M. J., and P. S. Cannon (2004), Assimilation of radio occultation measurements into background ionospheric models, edited by J. Goodman, *Radio Sci.*, 39(RS1S08), 100–106, doi:10.1029/2002RS002819.

Angling, M. J., and N. Jackson-Booth (2011), A short note on the assimilation of collocated and concurrent GPS and ionosonde data into the Electron Density Assimilative Model, *Radio Sci.*, 46, doi:10.1029/2010RS004566.

Angling, M. J., and B. Khattatov (2006), Comparative study of two assimilative models of the ionosphere, *Radio Sci.*, 41(RS5S20), doi:10.1029/2005RS003372.

Angling, M. J., J. Shaw, A. K. Shukla, and P. S. Cannon (2009), Development of an HF frequency selection tool based on the EDAM real time ionosphere, *Radio Sci.*, 44(RS0A13), doi:10.1029/2008RS004022.

Angling, M. J., P. S. Cannon, and P. Bradley (2013), Ionospheric propagation, in *Propagation of Radiowaves*, edited by L. W. Barclay, The Institution of Engineering and Technology, IET Radiowave propagation course.

Angrisano, A., S. Gaglione, C. Gioia, M. Massaro, and S. Troisi (2013), Benefit of the NeQuick Galileo Version in GNSS Single-Point Positioning, *Int. J. Navig. Obs.*, 2013, 1–11, doi:10.1155/2013/302947.

Appleton, E. V (1932), Wireless Studies of the Ionosphere, *J. Institution Electr. Eng.*, 71(430), 642–650.

- Augenbaum, J. M. (1984), A Lagrangian Method for the Shallow Water Equations Based on a Voronoi Mesh - One Dimensional Results, *J. Comput. Phys.*, *53*, 240–265.
- Barnston, A. G., S. J. Mason, L. Goddard, D. G. Dewitt, and S. E. Zebiak (2003), Multimodel Ensembling in Seasonal Climate Forecasting at IRI, *Bull. Am. Meteorol. Soc.*, *84*(12), 1783–1796, doi:10.1175/BAMS-84-12-1783.
- Belrose, J. S., and L. Thomas (1968), Ionization changes in the middle latitude D-region associated with geomagnetic storms, *J. Atmos. Terr. Phys.*, 1397–1413.
- Bilitza, D. (1990), *International Reference Ionosphere, 1990*, NASA, Greenbelt, Maryland, USA.
- Bilitza, D. (2004a), 35 years of International Reference Ionosphere - Karl Rawer's legacy, *Adv. Radio Sci.*, *2*, 283–287.
- Bilitza, D. (2004b), A correction for the IRI topside electron density model based on Alouette/ISIS topside sounder data, *Adv. Sp. Res.*, *33*(6), 838–843, doi:10.1016/j.asr.2003.07.009.
- Bilitza, D. (2006), The International Reference Ionosphere - Climatological Standard for the Ionosphere, in *Characterising the Ionosphere*, Neuilly-sur-Seine, France.
- Bilitza, D., and B. W. Reinisch (2008), International Reference Ionosphere 2007: Improvements and new parameters, *J. Adv. Sp. Res.*, *42*(4), 599–609, doi:10.1016/j.asr.2007.07.048.
- Bilitza, D., and R. Williamson (1999), Towards a Better Representation of the IRI Topside Based on ISIS and Alouette Data, *Adv. Sp. Res.*, *25*(1), 149–152.
- Bilitza, D., K. Rawer, and S. Pallaschke (1988), Study of Ionospheric Models for Satellite Orbit Determination, *Radio Sci.*, *23*(3), 223–232.
- Bott, A. (1989), A Positive Definite Advection Scheme Obtained by Nonlinear Renormalization of the Advective Fluxes, *Mon. Weather Rev.*, *117*, 1006–1015.
- Bruinsma, S. L., N. Sánchez-Ortiz, E. Olmedo, and N. Guijarro (2012), Evaluation of the DTM-2009 thermosphere model for benchmarking purposes, *J. Sp. Weather Sp. Clim.*, *2*, doi:10.1051/swsc/2012005.
- Bruyninx, C., M. Becker, and G. Stangl (2001), Regional densification of the IGS in Europe using the EUREF permanent GPS network (EPN), *Phys. Chem. Earth, Part A Solid Earth Geod.*, *26*(6-8), 531–538.
- Budden, K. G. (1985), *The propagation of radio waves. The theory of radio waves of low power in the ionosphere and magnetosphere*, Cambridge University Press.
- Buonsanto, M. J. (1999), Ionospheric Storms - A Review, *Space Sci. Rev.*, *88*, 563–601.

- Buresova, D., B. Nava, I. Galkin, M. J. Angling, S. M. Stankov, and P. Coisson (2009), Data ingestion and assimilation in ionospheric models, *Ann. Geophys.*, 52(3-4), 235–253.
- Bust, G. S., and C. N. Mitchell (2008), History, current state, and future directions of ionospheric imaging, *Rev. Geophys.*, 46(RG1003), 1–23, doi:10.1029/2006RG000212.
- Chandler, R. E. (2013), Exploiting strength, discounting weakness: combining information from multiple climate simulators, *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, 371.
- Chang, J. C., and S. R. Hanna (2004), Air quality model performance evaluation, *Meteorol. Atmos. Phys.*, 87, 167–196, doi:10.1007/s00703-003-0070-7.
- Chapman, S. (1931), The Absorption and Dissociative or Ionising Effect of Monochromatic Radiation in an Atmosphere on a Rotating Earth, *Proc. Phys. Soc.*, 43, 26.
- Chatfield, C. (1991), *The Analysis of Time Series*, 4th ed., Chapman and Hall.
- Codrescu, M., C. Negrea, M. Fedrizzi, T. Fuller-Rowell, A. Dobin, N. Jakowski, H. Khalsa, T. Matsuo, and N. Maruyama (2012), A Real-Time Run of the Coupled Thermosphere Ionosphere Plasmasphere Electrodynamics (CTIPE) Model, *Sp. Weather*, 10(2), doi:10.1029/2011SW000736.
- Coisson, P., B. Nava, and S. M. Radicella (2009), On the use of NeQuick topside option in IRI-2007, *Adv. Sp. Res.*, 43(11), 1688–1693, doi:10.1016/j.asr.2008.10.035.
- Crawford, F. S. (1968), *Waves (Berkeley Physics Course, Vol. 3)*, McGraw-Hill.
- Daniell, R. E., L. D. Brown, D. N. Anderson, M. W. Fox, P. H. Doherty, D. T. Decker, J. J. Sojka, and R. W. Schunk (1995), Parameterized ionospheric model: A global ionospheric parameterisation based on first principles models, *Radio Sci.*, 30(5), 1499.
- Davies, K. (1990), *Ionospheric Radio*, Peter Peregrinus Ltd., London.
- Dickey, D. A., and W. A. Fuller (1979), Distribution of the estimators for autoregressive time series with a unit root, *J. Am. Stat. Assoc.*, 74(366a), 427–431.
- Le Dimet, F.-X., and O. Talagrand (1986), Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects, *Tellus*, 38A, 97–110.
- Doblas-Reyes, F. J., M. Déqué, and J.-P. Pielikevire (2000), Multi-model spread and probabilistic seasonal forecasts in PROVOST, *Q. J. R. Meteorol. Soc.*, 126(567), 2069–2087.
- Doblas-Reyes, F. J., R. Hagedorn, and T. N. Palmer (2005), The rationale behind the success of multi-model ensembles in seasonal forecasting – II . Calibration and combination, *Tellus*, 2005(57A), 234–252.

- Dow, J. M., R. E. Neilan, and G. Gendt (2005), The International GPS Service: Celebrating the 10th anniversary and looking to the next decade, *Adv. Sp. Res.*, 36(3), 320–326, doi:10.1016/j.asr.2005.05.125.
- Dow, J. M., R. E. Neilan, and C. Rizos (2009), The International GNSS Service in a changing landscape of Global Navigation Satellite Systems, *J. Geod.*, 83(3-4), 191–198, doi:10.1007/s00190-008-0300-3.
- Dudeney, J. R. (1978), An improved model of electron concentration with height in the ionosphere, *J. Atmos. Terr. Phys.*, 40(2), 195–203.
- Eddy, A. (1967), The Statistical Objective Analysis of Scalar Data Fields, *J. Appl. Meteorol.*, 6, 597–609.
- Elvidge, S., and M. J. Angling (2014), Comparative testing of empirical, data assimilation and physics-based ionospheric/thermospheric models, in *URSI XXXI General Assembly*, Beijing.
- Elvidge, S., H. Godinez, M. J. Angling, and J. Koller (2013), *Improved modelling of upper atmospheric densities using multi-model ensembles*, edited by J. Koller and R. D. Gurule, Los Alamos National Laboratory.
- Elvidge, S., M. J. Angling, and B. Nava (2014), On the Use of Modified Taylor Diagrams to Compare Ionospheric Assimilation Models, *Radio Sci.*, doi:10.1002/2014RS005435.
- Emery, B. A., V. Coumans, D. S. Evans, G. A. Germany, M. S. Greer, E. Holeman, K. Kadinsky-Cade, F. J. Rich, and W. Xu (2008), Seasonal, Kp, solar wind, and solar flux variations in long-term single-pass satellite estimates of electron and ion auroral hemispheric power, *J. Geophys. Res.*, 113(A6), A06311, doi:10.1029/2007JA012866.
- Enders, W. (2009), *Applied Econometric Times Series*, 4th ed., John Wiley & Sons.
- Epstein, P. S. (1930), Reflection of Waves in an Inhomogeneous Absorbing Medium, *Proc. Natl. Acad. Sci.*, 16(10).
- Evans, R. E., M. S. J. Harrison, R. Graham, and K. R. Mylne (2000), Joint Medium-Range Ensembles from The Met Office and ECMWF Systems, *Mon. Weather Rev.*, 128(9), 3104–3127.
- Evensen, G. (1994), Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, 99(C5), 10143, doi:10.1029/94JC00572.
- Evensen, G. (2009), *Data Assimilation, the Ensemble Kalman Filter*, 2nd ed., Springer-Verlag Berlin Heidelberg.
- Ezquer, R. G., C. A. Jadur, and M. Mosert de Gonzalez (1998), IRI-95 TEC Predictions for the South American Peak of the Equatorial Anomaly, *Adv. Sp. Res.*, 22(6), 811–814.

- Fisher, R. A. (1915), Frequency Distribution of the Values of the Correlation Coefficient in Samples from an Indefinitely Large Population, *Biometrika*, 10(4), 507–521.
- Fisher, R. A. (1921), On the “probable error” of a coefficient of correlation deduced from a small sample, *Metron*, 1, 3–32.
- Francis, S. H. (1975), Global propagation of atmospheric gravity waves: A review, *J. Atmos. Terr. Phys.*, 37(6–7), 1011–1054.
- Freidenreich, S. M., S. T. Garner, R. G. Gudgel, C. Gordon, and I. M. Held (2004), The new GFDL global atmosphere and land model AM2-LM2: Evaluation with prescribed SST simulations", *J. Clim.*, 17(24), 4641–4673.
- Fridman, S. V, L. J. Nickisch, M. Aiello, and M. Hausman (2006), Real-time reconstruction of the three-dimensional ionosphere using data from a network of GPS receivers, *Radio Sci.*, 41(RS5S12), doi:10.1029/2005RS003341.
- Fridman, S. V, L. J. Nickisch, and M. Hausman (2009), Personal-computer-based system for real-time reconstruction of the three-dimensional ionosphere using data from diverse sources, *Radio Sci.*, 44(RS3008), doi:10.1029/2008RS004040.
- Fridman, S. V, L. J. Nickisch, and M. Hausman (2012), Inversion of backscattered ionograms and TEC data for over-the-horizon radar, *Radio Sci.*, 47, doi:10.1029/2011RS004932.
- Fritsch, J. M., J. Hilliker, and J. Ross (2000), Model Consensus, *Weather Forecast.*, 15(5), 571–582.
- Fuller-Rowell, T. J., and D. Rees (1981), A Three-Dimensional Time-Dependent Global Model of the Thermosphere, *J. Atmos. Sci.*, 37, 2545–2567.
- Fuller-Rowell, T. J., M. V. Codrescu, R. J. Moffett, and S. Quegan (1994), Response of the thermosphere and ionosphere to geomagnetic storms, *J. Geophys. Res.*, 99(A3), 3893–3914.
- Fuller-Rowell, T. J., M. V. Codrescu, and E. A. Araujo-Pradere (2001), Capturing the Storm-Time F-Region Ionospheric Response in an Empirical Model, in *Space Weather*, edited by P. Song, H. J. Singer, and G. L. Siscoe, American Geophysical Union, Washington, D. C.
- Galkin, I., and B. W. Reinisch (2008), The new ARTIST 5 for all digisondes, *Ina. Bull.*, 24.
- Gandin, L. (1963), *Objective analysis of meteorological fields (Leningrad: Gridromet). English translation (Jerusalem: Israel Program for Scientific Translation), 1965.*
- Giorgi, F., and L. O. Mearns (2002), Calculation of Average, Uncertainty Range, and Reliability of Regional Climate Changes from AOGCM Simulations via the “Reliability Ensemble Averaging” (REA) Method, *J. Clim.*, 15, 1141–1158.
- Gleckler, P. J., K. E. Taylor, and C. Doutriaux (2008), Performance metrics for climate models, *J. Geophys. Res.*, 113(D6), D06104, doi:10.1029/2007JD008972.

- Godinez, H. C. (2013), Personal Communication,
- Goodman, J. M., and J. Aarons (1990), Ionospheric effects on modern electronic systems, *Proc. IEEE*, 78(3), 512–528.
- Gorney, D. J. (1990), Solar cycle effects on the near-Earth space environment, *Rev. Geophys.*, 28, 315–336.
- Hagan, M. E., M. D. Burrage, J. M. Forbes, J. Hackney, W. J. Randel, and X. Zhang (1999), GSWM-98: Results for migrating solar tides, *J. Geophys. Res. Sp. Phys.*, 104(A4), 6813–6827, doi:10.1029/1998JA900125.
- Hagedorn, R., F. J. Doblas-Reyes, and T. N. Palmer (2005), The rationale behind the success of multi-model ensembles in seasonal forecasting – I. Basic concept, *Tellus*, 57A, 219–233.
- Hajj, G. A., B. D. Wilson, C. Wang, X. Pi, and G. Rosen (2004), Data assimilation of ground GPS total electron content into a physics-based ionospheric model by use of the Kalman filter, *Radio Sci.*, 39(RS1S05), doi:10.1029/2002RS002859.
- Hargreaves, J. K. (1979), *The upper atmosphere and solar-terrestrial relations*, Van Nostrand Reinhold Company.
- Harris, M. J. (2011), A New Coupled Middle Atmosphere and Thermosphere General Circulation Model: Studies of Dynamic, Energetic and Photochemical Coupling in the Middle and Upper Atmosphere, University College London.
- Harrison, M. S. J., T. N. Palmer, D. S. Richardson, R. Buizza, and T. Petroliaigis (1995), Joint ensembles from the UKMO and ECMWF models, in *ECMWF Seminar Proceedings: Predictability*, vol. 2, pp. 61–120.
- Hartree, D. R. (1929), The propagation of electromagnetic waves in a stratified medium, *Math. Proc. Cambridge Philos. Soc.*, 25(01), 97, doi:10.1017/S0305004100018600.
- Hedin, A. E. (1987), MSIS-86 thermospheric model, *J. Geophys. Res.*, 92, 4649–4662.
- Hedin, A. E. (1991), Extension of the MSIS thermospheric model into the middle and lower atmosphere, *J. Geophys. Res.*, 96, 1159–1172.
- Heelis, R. A., J. K. Lowell, and R. W. Spiro (1982), A model of the high-latitude ionospheric convection pattern, *J. Geophys. Res.*, 87, 6339–6345.
- Hocke, K., and K. Schlegel (1999), A review of atmospheric gravity waves and travelling ionospheric disturbances: 1982-1995, *Ann. Geophys.*, 14(9), 917–940, doi:10.1007/s00585-996-0917-6.
- Hofmann-Wellenhof, B., H. Lichtenegger, and J. Collins (2001), *Global positioning system: theory and practice*.

- Houtekamer, P. L., and H. L. Mitchell (2005), Ensemble Kalman Filtering, *Q. J. R. Meteorol. Soc.*, 131, 3269–3289, doi:10.1256/qj.05.135.
- Hunsucker, R. D. (1982), Atmospheric gravity waves generated in the high-latitude ionosphere: A review, *Rev. Geophys.*, 20(2), 293–315, doi:10.1029/RG020i002p00293.
- Hunsucker, R. D., and J. K. Hargreaves (2003), *The high-latitude ionosphere and its effects on radio propagation*, Cambridge University Press.
- IGS Central Bureau (2010), IGS Tracking Network, Available from: <http://igs.org/network/netindex.html> (Accessed 21 July 2014)
- ISO (2009), Space systems - Space environment (natural and artificial) - The Earth's ionosphere model: international reference ionosphere (IRI) model and extensions to the plasmasphere,
- Jacchia, L. G. (1977), Thermospheric temperature, density, and composition: new models, *SAO Spec. Rep.*, 375.
- Julier, S. J., and J. K. Uhlmann (2004), Unscented filtering and nonlinear estimation, *Proc. IEEE*, 92(3), 401–422.
- Justus, C. G., and D. L. Johnson (1999), *The NASA/MSFC Global Reference Atmosphere Model - 1999 Version (GRAM-99)*.
- Kalman, R. E. (1960), A New Approach to Linear Filtering and Prediction Problems, *Trans. ASME-Journal Basic Eng.*, 82, 34–45.
- Kalnay, E. (2003), *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge University Press, Cambridge.
- Kass, R. E., and A. E. Raftery (1996), Bayes factor, *J. Am. Stat. Assoc.*, 90.
- Kavanagh, A. J., S. R. Marple, F. Honary, I. W. McCrea, and A. Senior (2004), On solar protons and polar cap absorption: constraints on an empirical relationship, *Ann. Geophys.*, 22, 1133–1147.
- Kelley, M. C. (2009), *The Earth's Ionosphere: Plasma Physics & Electrodynamics*, International Geophysics, Elsevier Science.
- Kelly, A. (2013), *Mobile Robotics: Mathematics, Models, and Methods*, Cambridge University Press.
- Kenny, D. A. (1987), *Statistics for the Social and Behavioral Sciences*, Little, Brown.
- Kessler, D. J., N. L. Johnson, J. C. Liou, and M. Matney (2010), The Kessler syndrome: implications to future space operations, *Adv. Astronaut. Sci.*, 137(8), 2010.

- Kharin, V. V, and F. W. Zwiers (2002), Climate Predictions with Multimodel Ensembles, *J. Clim.*, *15*, 793–799.
- Knipp, D. J. (2011), *Understanding Space Weather and the Physics Behind It*, edited by M. McQuade and D. Kirkpatrick, McGraw-Hill.
- Kobayashi, C., H. Endo, Y. Ota, S. Kobayashi, H. Onoda, Y. Harada, K. Onogi, and H. Kamahori (2014), Preliminary Results of the JRA-55C, an Atmospheric Reanalysis Assimilating Conventional Observations Only, *Sola*, *10*, 78–82, doi:10.2151/sola.2014-016.
- Koller, J. (2012), Avoiding Collisions in Space, in *NNSA LDRD Symposium*, Washington DC.
- Krishnamurti, T. N. (1999), Improved Weather and Seasonal Climate Forecasts from Multimodel Superensemble, *Science* (80-.), *285*(5433), 1548–1550, doi:10.1126/science.285.5433.1548.
- Krishnamurti, T. N., C. M. Kishtawa, Z. Zhang, T. LaRow, D. Bachiochi, E. Williford, and S. Surendran (2000), Multimodel Ensemble Forecasts for Weather and Seasonal Climate, *J. Clim.*, *13*, 4196–4216.
- Lassen, H. (1927), Die taglichen Schwankungen des Ionisationszustandes der Heaviside-Schicht., *Elektrische Nachrichten-Technik*, *4*(4), 174–179.
- Lastovicka, J. (1996), Effects of geomagnetic storms in the lower ionosphere, middle atmosphere and troposphere, *J. Atmos. Terr. Phys.*, *58*(7), 831–843.
- Legendre, A. M. (1805), *Nouvelles méthodes pour la détermination des orbites des cometes*, F. Didot.
- Llewellyn, S. K., and R. B. Bent (1973), *Documentation and Description of the Bent Ionospheric Model*, Hanscom AFB, MA, USA.
- Lowell Digisonde International (2014a), Instrument Description, , (17/01/2012). Available from: <http://www.digisonde.com/instrument-description.html> (Accessed 21 July 2014)
- Lowell Digisonde International (2014b), Station Map, Available from: <http://www.digisonde.com/digisonde-station-map.html>
- Mandrake, L., B. Wilson, C. Wang, G. A. Hajj, A. J. Mannucci, and X. Pi (2005), A performance evaluation of the operational Jet Propulsion Laboratory/University of Southern California Global Assimilation Ionospheric Model (JPL/USC GAIM), *J. Geophys. Res.*, *110*(A12306), doi:doi:10.1029/2005JA011170.
- Maraun, D. et al. (2010), Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, *Rev. Geophys.*, *48*(3).

- Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen (2002), A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States, *J. Clim.*, *15*(22), 3237–3251.
- McLaughlin, C. A., A. Hiatt, and T. Lechtenberg (2011), Precision orbit derived total density, *J. Spacecr. Rockets*, *48*(1), 166–174.
- McNamara, L. F. (2009), Spatial correlations of foF2 deviations and their implications for global ionospheric models: 2. Digisondes in the United States, Europe, and South Africa, *Radio Sci.*, *44*, doi:10.1029/2008RS003956.
- McNamara, L. F. (2010), The Role of Ionosondes in Global Ionospheric Modeling, in *AGU Fall Meeting*.
- McNamara, L. F., C. R. Baker, and D. T. Decker (2008), Accuracy of USU-GAIM Specifications of foF2 and M(3000)F2 for a World-Wide Distribution of Ionosonde Locations, *Radio Sci.*, *43*(RS1011), doi:doi:10.1029/2007RS003754.
- McNamara, L. F., M. J. Angling, S. Elvidge, S. V. Fridman, M. A. Hausman, L. J. Nickisch, and L.-A. McKinnell (2013), Assimilation procedures for updating ionospheric profiles below the F2 peak, *Radio Sci.*, *48*(2), 143–157, doi:10.1002/rds.20020.
- Miller, R. L., G. A. Schmidt, and D. T. Shindell (2006), Forced annular variations in the 20th century Intergovernmental Panel on Climate Change Fourth Assessment Report models, *J. Geophys. Res.*, *111*, doi:10.1029/2005JD006323.
- Mitchell, C. (2013), The ionosphere, in *Propagation of Radiowaves*, edited by L. W. Barclay, The Institution of Engineering and Technology.
- Mitchell, C. N., and P. S. J. Spencer (2003), A three-dimensional time-dependent algorithm for ionospheric imaging using GPS, *Ann. Geophys.*, *46*(4), 687–696.
- Mote, P. W., and E. P. Salathe (2010), Future climate in the Pacific Northwest, *Clim. Chang.*, *102*, 29–50, doi:10.1007/s10584-010-9848-z.
- NASA Space Vehicle Criteria (Environment) (1973), *Models of Earth's Atmosphere (90 to 2500 km)*.
- National Space Weather Program (1995), *Strategic Plan*, Washington, DC.
- Nava, B., S. Radicella, R. Leitinger, and P. Coisson (2006), A near-real-time model-assisted ionosphere electron density retrieval method, *Radio Sci.*, *41*, doi:10.1029/2005RS003386.
- Nava, B., P. Coisson, and S. Radicella (2008), A new version of the neQuick ionosphere electron density model, *J. Atmos. Sol. Terr. Phys.*, *70*(15), 1856–1862, doi:10.1016/j.jatosp.2008.01.015.

- Ott, E., B. Hunt, I. Szunyogh, A. Zimin, E. Kostelich, M. Corazza, E. Kalnay, and J. Yorke (2004), A local ensemble Kalman filter for atmospheric data assimilation, *Tellus*, *56*, 415–428.
- Palmer, P. I., J. J. Barnett, J. R. Eyre, and S. B. Healy (2000), A non-linear optimal estimation inverse method for radio occultation measurements of temperature, humidity and surface pressure, *J. Geophys. Res. - Atmos.*, *105*(D13), 17513–17526.
- Palmer, T. N., and J. Shukla (2000), Editorial (for special issue on DSP/PROVOST), *Q. J. R. Meteorol. Soc.*, *126*, 1989–1990.
- Palmer, T. N. et al. (2004), Development of a European Multimodel Ensemble System for Seasonal-To-Interannual Prediction (DEMETER), *Bull. Am. Meteorol. Soc.*, *85*(6), 853–872, doi:10.1175/BAMS-85-6-853.
- Park, S. K., and L. Xu (2013), *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. II)*, SpringerLink : B{ü}cher, Springer.
- Parsons, A. M. (2006), Modeling E & F Region Ionospheric Response to X-Ray Solar Flares, 76 pp.
- Pavan, V., and F. J. Doblas-Reyes (2000), Multi-model seasonal hindcasts over the Euro-Atlantic: skill scores and dynamic features, *Clim. Dyn.*, *16*(8), 611–625, doi:10.1007/s003820000063.
- Picone, J. M., A. E. Hedin, D. P. Drob, and A. C. Aikin (2002), NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues, *J. Geophys. Res.*, *107*, 1468.
- Piggott, W. R., and K. Rawer (1978), *URSI Handbook of Ionogram Interpretation and Reduction*, 2nd ed.
- Plackett, R. L. (1950), Some theorems in least squares, *Biometrika*, *37*(1-2), 149–57.
- Purnell, D. K. (1976), Solution of the Advective Equation by Upstream Interpolation with a Cubic Spline, *Mon. Weather Rev.*, *104*, 42–48.
- Qian, L., S. C. Solomon, and T. J. Kane (2009), Seasonal variation of thermospheric density and composition, *J. Geophys. Res.*, *114*.
- Rajagopalan, B., U. Lall, and S. E. Zebiak (2002), Categorical Climate Forecasts through Regularization and Optimal Combination of Multiple GCM Ensembles, *Mon. Weather Rev.*, *130*, 1792–1811.
- Rastogi, R. G. (1958), Geomagnetic influence on the F1- and F2-regions of the ionosphere-effect of solar activity, *J. Atmos. Terr. Phys.*, *14*, 31–40.
- Ratcliffe, J. A. (1959), *The magneto-ionic theory and its applications to the ionosphere*, University Press Cambridge.

- Rawer, K. (1983), Replacement of the Present Sub-Peark Plasma Density Profile by a Unique Expression, *Adv. Sp. Res.*, 2(10), 183–190.
- Rawer, K., and K. Suchy (1976), Remarks concerning the dispersion equation of electromagnetic waves in a magnetized cold plasma, *J. Atmos. Terr. Phys.*, 38, 395–398.
- Reigber, C., H. Luhr, and P. Schwintzer (2002), CHAMP Mission Status, *Adv. Sp. Res.*, 30(2), 129–134.
- Reigber, C., H. Luhr, and P. Schwintzer (2003), First CHAMP mission results for gravity, magnetic and atmospheric studies,
- Reinisch, B. W. (1996), Modern ionosondes, in *Modern Ionospheric Science*, edited by H. Kohl, R. Ruster, and K. Schlegel, pp. 440–458, European Geophysical Society, Katlenburg-Lindau, Germany.
- Reinisch, B. W., and X. Huang (1983), Automatic calculation of electron density profiles from digital ionograms, 3, Processing of bottomside ionograms, *Radio Sci.*, 18, 477–492.
- Reinisch, B. W. et al. (2009), The New Digisonde for Research and Monitoring Applications, *Radio Sci.*, 44(RS0A24), doi:10.1029/2008RS004115.
- Richmond, A. D. (1992), Assimilative mapping of ionospheric electrodynamics, *Adv. Sp. Res.*, 12(6), 59–68.
- Richmond, A. D., E. C. Ridley, and R. G. Roble (1992), A thermosphere/ionosphere general circulation model with coupled electrodynamics, *Geophys. Res. Lett.*, 601–604.
- Ridley, A. J., Y. Deng, and G. Toth (2006), The Global Ionosphere-Thermosphere Model (GITM), *J. Atmos. Sol. Terr. Phys.*, 68, 839–864.
- Rishbeth, H., and P. J. S. Williams (1985), The EISCAT ionospheric radar - The system and its early results, *R. Astron. Soc. Q. J.*, 26, 478–512.
- Rissanen, J. (1983), A universal prior for integers and estimation by minimum description length, *Ann. Stat.*, 11, 416–431.
- Roble, R. G., E. C. Ridley, A. D. Richmond, and R. E. Dickinson (1988), A coupled thermosphere/ionosphere general circulation model, *Geophys. Res. Lett.*, 15(12), 1325–1328.
- Rodgers, C. D. (2000), *Inverse methods for atmospheric sounding: theory and practice*, *Atmospheric, Oceanic and Planetary Physics*, World Scientific Publishing, Singapore.
- Schunk, R. W. (1996), *Solar Terrestrial Energy Program (STEP). Handbook of Ionospheric Models*, SCOSTEP, Boulder, Colorado, USA.
- Schwarz, G. (1978), Estimating the dimension of a model, *Ann. Stat.*, 6, 461–464.

Sheskin, D. J. (2004), *Handbook of parametric and nonparametric statistical procedures*, edited by C. H. Crc, Chapman & Hall/CRC.

Shim, J. S. et al. (2011), CEDAR Electrodynamics Thermosphere Ionosphere (ETI) Challenge for systematic assessment of ionosphere/thermosphere models: NmF2, hmF2, and vertical drift using ground-based observations, *Sp. Weather*, 9, doi:10.1029/2011SW000727.

Shim, J. S. et al. (2012), CEDAR Electrodynamics Thermosphere Ionosphere (ETI) Challenge for systematic assessment of ionosphere/thermosphere models: Electron density, neutral density, NmF2, and hmF2 using space based observations, *Sp. Weather*, 10, doi:10.1029/2012SW000851.

Solomon, S. C. (2006), Numerical models of the E-region ionosphere, *Adv. Sp. Res.*, 37(5), 1031–1037, doi:10.1016/j.asr.2005.09.040.

Space Weather Prediction Center (2009), Glossary of solar-Terrestrial Terms, Available from: <http://www.swpc.noaa.gov/info/glossary.html> (Accessed 22 July 2014)

Spencer, P. S. J., and C. N. Mitchell (2007), Imaging of fast moving electron-density structures in the polar cap, *Ann. Geophys.*, 50(3), 427–434.

Stone, M. (1974), Cross-validators choice and assessment of statistical predictions (with Discussion), *J. R. Stat. Soc., Series B*(36), 111–147.

Sunil Rao, J. (2000), Bootstrapping to Assess and Improve Atmospheric Prediction Models, in *Statistical Mining and Data Visualization in Atmospheric Sciences*, vol. 41, edited by T. J. Brown and P. W. Mielke Jr, pp. 29–41, Springer US.

Sutton, E. K. (2009), Normalized Force Coefficients for Satellites with Elongated Shapes, *J. Spacecr. Rockets*, 46(1), 112–116.

Sutton, E. K., J. M. Forbes, and R. S. Nerem (2005), Global thermospheric neutral density and wind response to the severe 2003 geomagnetic storms from CHAMP accelerometer data, *J. Geophys. Res.*, 110(A9), A09S40, doi:10.1029/2004JA010985.

Taylor, K. E. (2001), Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.*, 106(7), 7183–7192.

Tebaldi, C., and R. Knutti (2007), The use of the multi-model ensemble in probabilistic climate projections, *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, 365(1857), 2053–2075, doi:10.1098/rsta.2007.2076.

Thompson, P. D. (1977), How to improve Accuracy by Combining Independent Forecasts, *Mon. Weather Rev.*, 105(2), 228–229.

Tikhonov, A. N., and V. Y. Arsenin (1977), *Solution of Ill-Posed Problems*, Halsted, New York.

- Titheridge, J. (1988), The real height analysis of ionograms: A generalized formulation, *Radio Sci.*, 23(5), 831–849.
- Tracton, S., and E. Kalnay (1993), Operational Ensemble Prediction at the National Meteorological Center: Practical Aspects, *Weather Forecast.*, 8, 379–398.
- Triskova, L., V. Truhlik, J. Smilauer, and C. Republic (2002), On Possible Improvements of Outer Ionosphere Ion Composition Model in IRI, *Adv. Sp. Res.*, 29(6), 849–858.
- U.S. Dept. of Commerce, NOAA, S. W. P. C. (2007), NOAA POES Hemispheric Power Information, Available from: <http://www.swpc.noaa.gov/ftpdir/lists/hpi/README.txt>
- Vallado, D. A., and D. Finkleman (2008), A critical assessment of satellite drag and atmospheric density modeling, *Astrodyn. Spec. Conf. Exhib.*, 18–21.
- Vislocky, R. L., and J. M. Fritsch (1995), Improved Model Output Statistics Forecasts through Model Consensus, *Bull. Am. Meteorol. Soc.*, 76(7), 1157–1164.
- Weimer, D. R. (2005), Predicting surface geomagnetic variations using ionospheric electrodynamic models, *J. Geophys. Res. A Sp. Phys.*, 110(A12).
- Weisheimer, A. F., F. J. Doblas-Reyes, T. N. Palmer, A. Alessandri, A. Arribas, M. Déqué, and P. Roger (2009), Ensembles: A new multi-model ensemble for seasonal-to-annual predictions - Skill and progress beyond DEMETER in forecasting tropical Pacific SSTs, *Geophys. Res. Lett.*, 36.
- Wright, J. R. (2003), *Real-Time Estimation of Local Atmospheric Density*.
- Xu, G. (2007), *GPS: Theory, Algorithms and Applications*, Springer.
- Yau, A., T. Abe, and B. Whalen (1996), Cold plasma source of upflowing ionospheric ions in the nightside auroral ionosphere, *J. Geomagn. Geoelectr.*, (1972), 947–957.