THE UNIVERSITY
of ADELAIDE

FACULTY OF SCIENCES
SCHOOL OF PHYSICAL SCIENCES

A new empirical climatological model of ionospheric foF2 and hmF2 and review of the International Reference Ionosphere

Daniel Field

Supervisors:
Dr Andrew MacKinnon
Dr Trevor Harris
Dr Manuel Cervera

October 2018

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give permission for the digital version of my thesis to be made available on the web, via the Universitys digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Signature

29/10/18
Date

Abstract

This thesis presents Mocha, a new empirical climatological monthly median and variance model of two ionospheric parameters viz. foF2 and hmF2. Understanding ionospheric variance is important for assessing the climatological performance of systems which use trans-ionospheric paths such as satellite navigation systems, or utilise the refractive properties of the ionosphere to propagate high frequency radio waves long distances such as over-the-horizon radar (OTHR). Additionally, this thesis provides a global review of the International Reference Ionosphere 2016 (IRI2016) foF2 and hm2 models. IRI is an widely used, industry standard, global empirical climatological monthly median model of the ionosphere. Mocha produces slightly more accurate foF2 predictions and significantly more accurate hmF2 predictions than IRI. It also outputs the variance in these parameters which is not currently produced by IRI. Mocha is intended to be used in conjunction with IRI: its foF2 and hmF2 output parameters may be input directly into IRI overriding IRI's internal model of these parameters.

Acknowledgements

I am extremely thankful to the hard work of my supervisors Andrew MacKinnon, Manuel Cervera, Trevor Harris, David Holdsworth and David Neudegg, who have provided an enormous amount of work and effort to get this thesis up to a suitable level. They have provided great guidance and support over the last two years, which helped me dramatically improve the quality of my presentations and results.

I would also like to thank Baden, Tom, Lenard, Simon, Andrew H and Andrew S for providing an enjoyable degree filled with office antics. Finally, I would like to thank my family, Brayden, Kirsty and Adam for helping me throughout my degree.

We are grateful to the World Data Centre of the Australian Bureau of Meteorology, Space Weather Services for the provision of foF2 data.

We acknowledge using data of the Global Ionosphere Radio Observatory (GIRO), http://giro.uml.edu/.

We are grateful to the Defence Science and Technology Group of the Australian Federal Government for the provision of foF2 and hmF2 data.

Contents

Chapter 1

Introduction

1.1 Motivation

Understanding the ionosphere is important for determining its effects on instruments which employ trans-ionospheric paths such as satellite navigational systems (e.g. Global Positioning System (GPS)) or satellite communications. Radio frequency (RF) systems such as high frequency (HF) radio communications and over-the-horizon radar (OTHR) utilise the refractory properties of the ionosphere to propagate HF radio waves long distances beyond the horizon. Such HF systems have a wide array of uses for both defence and civilian purposes. As the ionosphere is the propagating environment which enables these technologies, it clearly plays a crucial role in how they operate and perform. The ionospheric F2 layer is the most important for the operation of HF communication systems and OTH radar.

Models of the ionosphere are important tools which enable users to estimate how HF systems are affected by it. There are many levels of models from climatology at one end of the model spectrum to models of space weather effects at the other end. Each of these models is important in different contexts. For example, climatological models of the ionosphere are important to assess how well HF systems such as OTH radar perform over the broad range of conditions (diurnal, seasonal and solar cycle) in which they are expected to operate.

This thesis presents Mocha, a new empirical climatological monthly median and variance model of two ionospheric F2 parameters viz. foF2 and hmF2. Additionally, this thesis provides a global review of the International Reference Ionosphere 2016 (IRI2016) foF2 and hmF2 models. IRI is a widely used, industry standard, global empirical climatological monthly median model of the ionosphere. Mocha produces slightly more accurate foF2 predictions and significantly more accurate hmF2 predictions than IRI. It also outputs the variance in these parameters which is not currently produced by IRI. Mocha is intended to be used in conjunction with IRI: its foF2 and hmF2 output parameters may be input directly into IRI overriding IRI's internal model of these parameters.

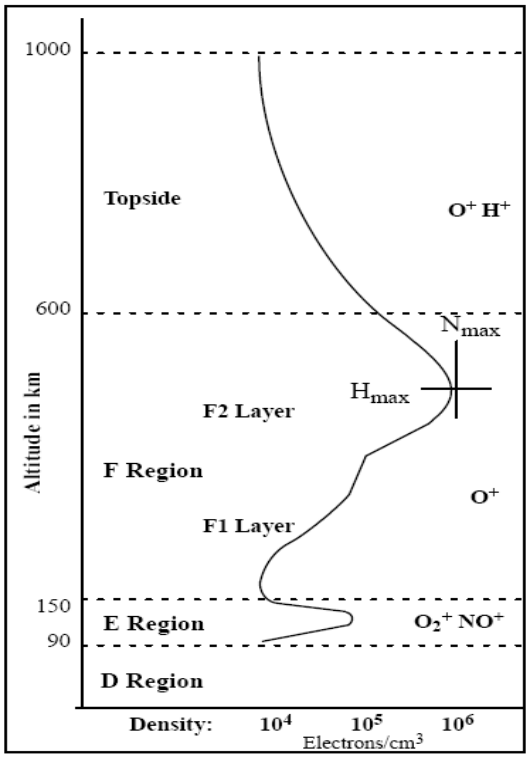

Figure 1.1: The different ionospheric layers, Dave Anderson and Tim Fuller-Rowell, 1999, *IonosphereProfileNOAA*, Accessed 6/10/2018, from https://commons.wikimedia.org/wiki/File:IonosphereProfileNOAA.png

1.2 The Ionosphere

The ionosphere is the region of the atmosphere which has been ionised by solar radiation and is capable of affecting the propagation of radio waves via refraction. The ionosphere is composed of three main regions, D, E and F, with heights ranging from 50 km to 90 km, 90 km to 140 km and above 140 km respectively. Each of these regions is composed of sublayers. The D region contains the D and C layers (Davies, 1990, p. 1). The E region contains the E layer and occasionally a sporadic layer, Es, which may be highly ionised or sometimes patchy in nature. Finally, the F region contains two sub layers F1 and F2. An example image of the ionospheric layers is displayed in Figure 1.1.

Each of these layers are produced via different mechanisms (Davies, 1990, p. 61). The D layer is ionised by hard x-rays and from the Lyman-alpha line while the C layer is produced by cosmic rays. The E layer is produced by solar soft x-rays and the F layer is ionised by solar Extreme Ultra Violet radiation (EUV). The D, E and F1 layers are solar controlled and are only present during the day, while the F2 layer is present at night but at a reduced electron density. To first order the F2 layer is solar controlled, but other factors such as ion transport, recombination rates and the Earth's magnetic field are important (IPS, 1994). All of these layers have

a unique critical frequency and height; e.g. foE and hmE for the E layer, and foF1 hmF1 for the F1 layer.

Radio waves are extremely useful for communication over large distances when satellites can not be used. The ionosphere allows long range propagation of HF signals via refraction. OTH radar are another HF system which uses the ionosphere, in this case for wide area surveillance for monitoring aircraft and maritime vessels. While microwave radar can be used to detect and track targets, they are limited to line of sight. This limitation is not present in OTHRs which are able to detect targets at much longer ranges. A typical OTH radar surveillance region is approximately a 90 ° arc from 1000 km to 3000 km (Fabrizio, 2013).

The ionosphere is birefringent at HF due to the geomagnetic field. HF radio waves upon entering the ionosphere will split into two canonical propagation modes viz. the ordinary (O) and extraordinary (X) modes. A vertically propagating O mode radio wave will be reflected at an altitude where the wave frequency is equal to the plasma frequency of the ionosphere. If the frequency of the radio wave is greater than the critical frequency of the F2 layer (foF2) the radio wave will penetrate the ionosphere and will not be reflected back to the ground. The altitude of the critical frequency of the F2 layer referred to as hmF2. Radio waves at progressively lower frequencies than the critical frequency will be reflected at progressively lower heights. foF2 and hmF2 have values ranging up to ~ 30 MHz and 450 km but these heights can be higher at the magnetic equator. foF2 and hmF2 are highly dependant on solar cycle, time of day, season and magnetic latitude.

Earth has a magnetic field that is produced from convection currents within the metallic outer core. On large scales Earth's magnetic field can be assumed to first order to act like a bar magnet tilted at ~ 11 ° to Earth's rotational axis. The orientation of the magnetic field can be characterised using the magnetic dip latitude and magnetic declination. The magnetic dip latitude is the angle between the magnetic field lines and Earth's surface. This parameter is important for describing the ionosphere, and is used to divide the ionosphere into three regions, low, mid and high magnetic latitudes. In general high magnetic latitudes are defined as | geomagnetic dip $| > 60$ °, mid latitudes is described by 60 ° > | geomagnetic dip | > 20 ° and finally low latitudes is given by | geomagnetic dip | < 20 °.

The ionosphere can be difficult to model as it can have an extremely complicated structure due to a variety of interacting effects: the neutral winds drive the plasma; the charged particles moving in electric ($\vec{E}$) and magnetic ($\vec{B}$) fields experience the Lorentz force $\vec{E} \times \vec{B}$. Plasma instabilities can cause a spread in foF2 observations and seasonal changes affect ion transport.

Plasma in the equatorial F region can exhibit instabilities under certain conditions causing equatorial spread F in the HF signal returns (Davies, 1990, p. 146). Measurements of the critical frequency show irregularities that can be spread in range or frequency. This phenomena typically only occurs at night though occasionally occurs during the day near the geomagnetic equator (Davies, 1990, p. 150).

Post sunset electron density inversions develop in the ionosphere as the lower layers recombine and disappear first. This can result in a Rayleigh-Taylor like instability at the geomagnetic equator. Small scale perturbations at the interface between these two layers can cause the collection of charge upon the edges of the perturbation. These charges give rise to then develop an electric field parallel to the two density layers. This perturbation electric field together with Earth's magnetic field causes an upwards plasma drift proportional to the cross product between the perturbation electric field and Earth's magnetic field. The lower density plasma is advected upwards and conversely the denser plasma is driven downwards. This drives larger perturbations which drives further plasma advection. This system is unstable as it is a non correcting feedback loop. The generalised Rayleigh-Taylor process (not described in detail here) allows for other forms of destabilising influences beyond gravity such as neutral winds (see e.g. Schunk and Nagy, 2009).

The electron density of the ionosphere maximises at the latitudes $\sim \pm 20\,^{\circ}$ from the magnetic equator and is referred to as the equatorial anomaly, (Davies, 1990, p. 124). This anomaly follows the Sun from East to West with a temporal lead behind the subsolar point (the point directly underneath the Sun). The equatorial anomaly is produced from the so called fountain effect whereby an upwards plasma drift is produced via interaction between Earth's magnetic field and neutral wind currents. A west to east electric field caused from the neutral winds combines with Earth's magnetic field driving the plasma upwards (i.e. the Lorentz force). This depletes the electron densities near the magnetic equator. The electrons then follow the magnetic field lines where they are preferentially diffused at $\pm$ 30 $^{\circ}$ geomagnetic dip creating two peaks in electron density North and South of the geomagnetic equator. Characterising this peak accurately requires a large number of ionospheric sounders at low latitudes.

The winter anomaly is a significant ionospheric feature that occurs at mid latitudes. Examination of the mid latitude ionosphere has shown that the winter foF2 values are twice that of the summer values, meaning the winter peak electron density was a factor of 4 times higher than in summer (Davies, 1990, p. 136). It was discovered that the winter anomaly effect occurs during day time only. During summer the foF2 value shows little diurnal variation unlike during winter. The decrease in the critical frequency in summer is believed to be caused by an increase in the molecular to atomic composition in the neutral atmosphere. This effect is known to be stronger in the Northern Hemisphere than in the Southern Hemisphere. Charac-

terising this effect accurately would require a large amount of mid latitude sounders that are regularly sampled throughout the year.

Ionospheric climatology is primarily influenced by the solar wind, a plasma of solar protons, electrons and alpha particles, which are emitted by the Sun. This material travels at $\sim 400\,\mathrm{km/s}$ and takes ~ 4 days to reach Earth. Earth is protected from the solar wind by the magnetosphere, a region where Earth's magnetic field is stronger than the interplanetary magnetic field. A shock wave between the solar wind and the magnetosphere, causes plasma deceleration and deflects the plasma around Earth. If the geomagnetic field lines are oppositely directed to the solar field lines then the solar plasma is directed towards Earth's geomagnetic poles producing auroras. The Sun is also responsible for space weather effects such as geomagnetic and ionospheric storms.

Solar events have a large effect on modern life affecting communications and power lines etc. (IPS, 1994). Solar radio bursts can affect satellite communications which use very high frequencies (VHF) and ultra high frequencies (UHF). This interference occurs predominately within the sunlit hemisphere of the Earth, around the equinoxes. There can also be effects on the low frequency (LF) and very low frequency (VLF) communications systems, as sudden changes to the ionosphere can cause phase anomalies in the signal (Kaufmann and Mendes, 1970). There are also effects on GPS, e.g. large solar bursts can cause a reduced signal to noise ratio which may result in GPS receivers losing lock on affected satellites resulting in reduced position accuracy or complete loss of position solution.

The flow of charged particles from the Sun can interfere with sensitive electronics in satellite systems (IPS, 1994). As charged particles build up on the body or deep in the dielectric components of a satellite, electrical discharge can then cause phantom commands which disrupt the control of the satellite and possibly lead to its loss. These effects can be dangerous for crewed spacecraft as astronauts could potentially experience a harmful amount solar radiation.

Magnetic storms can also result in damage to pipe and power lines. Storms can induce currents in buried conductors such as in pipelines, which can lead to dissimilar metal corrosion and reduction in lifetime (IPS, 1994). While methods have been developed to reduce this effect there can still be damage. Similar currents induced in power lines by geomagnetic storms can damage high voltage transformers and may result in the failure of a power grid. Power grids are expected to be more susceptible at higher latitudes where the effects of the geomagnetic storms are greater, but effects have been observed at mid latitudes.

Solar events such as coronal mass ejections leading to geomagnetic storms, can also effect mineral surveys (IPS, 1994). To search for minerals it is possible to use

ground-based or airborne magnetometers to measure changes in the geomagnetic field due to ore bodies. Geomagnetic storms can have major effects on this surveying technique, which can be rendered ineffective for the duration of the storm.

1.3 Observing Solar Activity

Solar forcing is primarily responsible for the condition of the ionosphere. Complicated solar features such as solar storms cause dramatic changes in Earth's ionosphere. As such understanding the Sun is required to understand the ionosphere.

The visible disk surrounding the Sun, the photosphere, emits ultraviolet (UV), infrared (IR) and visible light. It appears to have a texture like rice grains due to granulation which is caused by rising convective cells. This layer is roughly 200 km thick where the base temperature starts at about 6400 K but decreases steadily until it reaches about 4400 K.

Dark spots on the photosphere are known as sunspots. The sunspot number, R, is commonly used to indicate solar activity, (Davies, 1990, p. 29). R is a combination of the number of visible sunspot groups, g, and the number of individual sunspots f, given by,

$$R = k(f + 10g) \tag{1.1}$$

where k is a scaling factor based on each solar telescope. The most commonly used sunspot number index is R_{12}, which is a 12 month smoothed sunspot number (Davies, 1990, p. 31).

The sunspot number periodically cycles over an 11 year period, typical solar maxima have an $R \sim 150$. The current cycle is the 24th cycle since observations started in 1755. Throughout the solar cycle the magnetic structure of the Sun changes. At the start of a cycle, the north pole of the Sun may have a particular magnetic polarity, while the south pole will have the opposite polarity. During the middle of the cycle there will be no well defined polarity at either pole, and at the end of the cycle north and south poles switch polarity. At the start of the solar cycle sunspots form at latitudes around 30 to 35°. As the cycle progresses the location of where these sunspots first appear shifts towards the equator. Sunspots also have a magnetic polarity and are labelled as either a leader or a follower. A leader sunspot is the first spot in a series and has the same polarity as the hemisphere it is on, the trailer spots lag behind the leader and have an opposite polarity to the hemisphere.

Recently, there has been a change to the weighting of the sunspot number observations as described by Clette et al. (2014) and Clette and Lefèvre (2016). This change was to made to reduce inconsistencies between the sunspot number and group number measurements. The twelve month running mean sunspot number from version 1 observations is hearby referred to as R_{12}, and version 2 is referred to

as S_{12}. As discussed in these papers, version 2 removed the 0.6 Wolfer scaling factor and the Waldmeier weighting scheme which provided an approximate 1.17 average increase in sunspot number. This means that $R_{12} \approx (0.6 \times 1.17) S_{12} \approx (0.7) S_{12}$. This 0.7 scaling factor is used throughout the rest of this analysis to convert R_{12} to S_{12}.

Another way of characterising solar activity is to use the 10 cm radio flux, F10.7. Measurement of the emissions of the Sun at a wavelength of 10 cm are averaged over a period of a month or one year.

Short duration solar events effects such as solar flares and coronal mass ejections (CMEs) cause rapid ionospheric changes. Solar flares are a sudden release of energy in the Sun around a group of magnetically complicated or growing sunspot group (Davies, 1990, p. 313). CMEs form when a closed magnetic loop breaks and releases a large amount of energy, which forces plasma to leave the surface of the Sun, (Davies, 1990, p. 315). Filaments appear as dark thin structure on the Sun's disk, which can erupt to produce a CME (Davies, 1990, p. 315). As the analysis described in later chapters is focused on ionospheric climatology, these events are removed from the observations as discussed in Chapter 3.

1.4 Measurement Tools

Ionospheric properties are measured with a variety of instruments. Each instrument has its own strengths and weaknesses. These instruments make measurements either remotely, such as ionosondes, or directly through in situ rocket measurement. As this analysis only made use of ionosonde measurements this subsection only discusses ionosondes in detail.

Ionosondes typically sweep in radio frequency from $\sim 2\,\mathrm{MHz}$ to $\sim 30\,\mathrm{MHz}$ to characterise the ionosphere at different wavelengths (Davies, 1990, p. 89). Transmissions from a vertical incidence sounder will be reflected back to the ground at heights where the wave frequency is equal to the plasma frequency. Ground based receivers measure the incoming signal to determine ionospheric parameters such as the plasma frequency and virtual height.

The simplest measurement from a vertically incident ionosonde is the time of flight, T, which corresponds to a virtual height, the height that a radio wave travelling vertically at the speed of light in a vacuum would reach in half of the time of flight. The real or true height of the reflection is less than the virtual height due to the group speed of the radio waves being retarded by the ionosphere. The time of flight can also be determined by measuring the phase, ϕ:

$$2\pi T = \frac{d\phi}{df}. \tag{1.2}$$

Ionosondes typically have a phase accuracy of $\sim 4^\circ$ and a Δf of $\sim$ kHz giving a virtual height accuracy of ~ 0.2 km.

After an ionogram is produced, it must be scaled to identify ionospheric parameters such as foF2 and hmF2. There are two types of scaling, manual and automatic (Piggott and Rawer, 1978). Manual scaling involves a trained human scaler examining each ionogram and selecting critical parameters while automatic scaling is the same procedure, but using software. Guidelines for scaling algorithms are defined in (Piggott and Rawer, 1978). Manual scaling is generally regarded as more accurate as humans are more capable of dealing with the wide range of disturbed ionograms which may be recorded. However, manual scaling is human resource intensive and therefore expensive and slow compared to automatic scaling.

To accurately autoscale ionograms it is necessary to be able to distinguish and separate the ordinary (O) and extraordinary (X) mode returns. Near the magnetic dip equator, both the ordinary and extraordinary modes are linearly polarised with polarisation vectors orientated North-South (O) and East-West (X). Separation of the two modes is straight forward using two orthogonal antennas aligned in these directions. Separation of the two modes is more complicated in mid and high latitude regions. Unlike the low latitude regime where the O and X modes are in orthogonal linear orientations, at the mid to high latitude regions these modes have counter rotating elliptical polarisation vectors. Each antenna will receive a mixture of the two modes. If a ninety degree phase shift is added into one of the signals relative to the signal then using algebraic manipulation is possible to separate these two modes (Davies, 1990, p. 92). Recently more robust techniques using digital HF receivers have been developed (Harris and Pederick, 2017).

While ionosondes are ubiquitous as an ionospheric measurement instrument they do have some limitations. The large frequency range over which they operate, means that it can be difficult to optimise the performance for all frequencies. There are also other issues, for example solar flares can increase the strength of D region leading to strong absorption of the HF signal. This effect is referred to as a short wave fade out. Strong blanketing Es layers can obscure the F1 and F2 layers. Radio frequency interference (RFI) can make ionogram interpretation more difficult, particularly for automated scaling systems. Travelling ionospheric disturbances (TIDs) can affect the interpretation of the ionograms by distortion of the layer trace.

1.5 Selected sites for detailed analysis

To examine the global structure of IRI, the analysis later in this thesis focussed on several selected sites. Figures 1.2 and 1.3 display these selected sites for the world and Australia, along with lines indicating the geomagnetic equator and $\pm 30^\circ$ dip

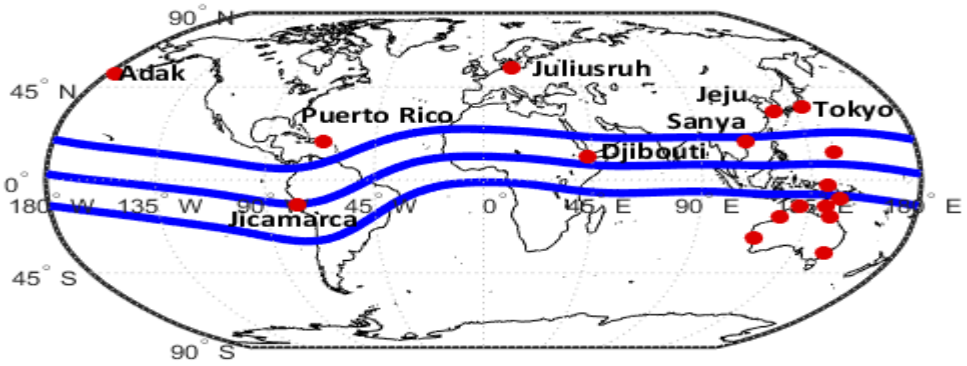

Figure 1.2: A global map that highlights selected sites from this analysis in red. It also includes the geomagnetic equator and $\pm 30\,^{\circ}$ geomagnetic dip angle's (blue lines).

angles, near where the equatorial ionospheric anomalies are expected to peak.

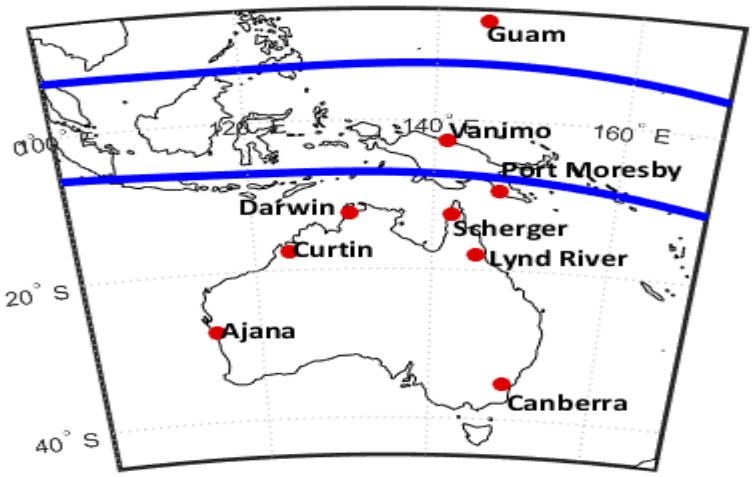

Figure 1.3: An Australasian map that highlights selected sites from this analysis in red. It also includes the geomagnetic equator (top blue line) and - 30 ° geomagnetic dip (lower blue line).

1.6 Summary

The ionosphere has important uses for military and civilian life and many radio systems are either affected by it (e.g. GPS) or rely on it for their operation (e.g. HF communications, OTHR). The operation of these systems can be improved using ionospheric models to assess their performance over a broad range of conditions. The International Reference Ionosphere model currently does not provide variance of the predictions of ionospheric parameters. The analysis in later chapters will produce a climatological monthly median model of foF2 and hmF2 along with the corresponding variance.

The analysis described in subsequent chapters uses a mixture of automatic and manually scaled foF2 and hmF2 observations to test IRI and develop a new model. In Chapter 2, the different ionospheric modelling techniques and types will be discussed and then the history and description of IRI is presented. Chapter 3 discusses the datasets used for the analysis and details the various filtering processes applied to the data. Chapter 4 provides an assessment of IRI by examining the seasonal, solar cycle and spatial and temporal errors. Finally, Chapter 5 describes various methodologies that were used to develop a new foF2 and hmF2 model. The new model is then analysed for seasonal, solar cycle and spatial influence on errors and these results are then compared to IRI's foF2 and hmF2 median model.

Chapter 2

International Reference Ionosphere (IRI) Overview

2.1 Modelling Techniques

There are two main methods for ionospheric prediction: empirical and physical models. Empirical models rely on a large collection of historical data, typically sampled sporadically in space and time. Physics based models use derived relations for ionisation rates, ion re-combination and ion transport by neutral winds and/or geomagnetism to provide predictions. Both models have their advantages and disadvantages. Some physical principles models include the *Sheffield University Plasmasphere Ionosphere Model* (SUPIM) (Bailey et al., 1997), the *Global Assimilation of Ionospheric Measurements - Full Physics Model* (GAIM-FP) (Schunk et al., 2004), NCAR's *Thermosphere and Ionosphere Electrodynamics Global Circulation Model*, TIEGCM (Richmond et al., 1992) and Naval Research Laboratory's (NRL) *SAMI3* (Huba et al., 2008) . The most commonly used empirical model is the *International Reference Ionosphere* (IRI) (Bilitza et al., 2017), a joint project between the Committee on Space Research (COSPAR) and the International Union of Radio Science (URSI).

Empirical models are solely based upon observations and consequently are simpler than physics models; they do not require a detailed understanding of the physical processes which occur in the ionosphere. They have several advantages such as relatively low computational requirements. They are capable of predicting ionospheric parameters based only upon the geographic coordinates, time and expected sunspot number, and can easily accommodate additional or missing observations and include new observations into their predictions. However, empirical models require a full solar cycle of data to be accurate and they tend to require high spatial sampling.

Physics based models obtain numerical solutions to conservation equations for energy, momentum, etc. for ions and electrons to calculate plasma densities, temperatures and flow velocities. Unfortunately, these models require magnetospheric and atmospheric inputs which can be difficult to determine, and thus require other models to define these required boundary conditions. This can increase the model

uncertainty making them less accurate. Empirical models do not have this problem as they are self contained. Unlike empirical models, physics models do allow investigation of the sensitivity of the ionospheric parameters to a wide range of different conditions.

Climatological models are simpler than real time models as they are used to obtain estimates of the background conditions for large spatial structures with long time scales. However, the ionosphere can vary dramatically from the median values due to events such as geomagnetic storms. Storms are capable of significantly affecting the ionosphere for very short time-scales, which a monthly median model would be incapable of representing. Furthermore, spatial features less than the typical sampling size site separation are not reproduced in these models. Overall, these models are useful for a broad understanding of the ionosphere for storm free conditions.

In comparison real time models, while being typically more computationally intensive, are capable of making predictions of the present ionosphere (nowcasting) and producing accurate predictions for up to hours into the future. Real time models are extremely useful for predicting HF propagation, which is used in radio communications, direction finding (HFDF) and OTHR. Empirical real time models have difficulties in producing rapid forecasts due to delayed receiving of observations and processing time in the model.

2.2 International Reference Ionosphere

The International Reference Ionosphere (IRI) (Bilitza et al., 2017) is a standard empirical model used to predict ionospheric properties such as foF2 and hmF2. IRI is a useful tool for looking at the general properties of the ionosphere in regions without instrumentation. IRI is used to predict future values at a given location and time of year, as only the sunspot number is required. IRI is capable of showing some detail of both the large and complex equatorial anomaly regions and the more stable mid latitude regions. However, it is better optimised for the Northern Hemisphere due to the higher density of observation sites in the NH.

In 1961 initial development on ionospheric models was started by the Committee on Space Research (COSPAR) who developed the COSPAR International Reference Atmosphere or CIRA for short (Bilitza, 2004). CIRA was initially developed for the purpose of the development of space based instruments, satellite altimetry data and radio astronomy. This model produced details about the thermosphere and ionosphere but a more accurate model of the ionosphere was required, and in 1968 the IRI project was initiated and chaired by Karl Rawer. A year later the International Union of Radio Science (URSI) joined with COSPAR to develop an ionospheric model but each with a different focus. URSI were interested in the electron den-

sity parameters which are important for understanding radio wave propagation. In comparison, COSPAR were more interested in the ionosphere's general properties for examining the environmental effects on spacecraft. In 1972 the first set of preliminary data was released for parameters at selected locations. In the following year at a COSPAR symposium the guidelines for input data in D-region modelling was established. In 1977 the global coverage of electron densities for foE, foF1, foF2 and M(3000)F2 were released. M(3000)F2 is defined as MUF(3000)/foF2, where MUF(3000) is the maximum usable frequency which can ionospherically propagate to a point on the ground 3000 km away. In 1995 IRI-95 was released which had improvements at low magnetic latitudes. The rise of the internet resulted in IRIWEB so IRI plots and data could be accessed online. When IRI-2001 was released it improved results for the D and F1 region, and added parameters for the F1 probability and equatorial vertical ion drift (Bilitza, 2001). A few years later IRI-2007 added a new approach for modelling the sodium topside along with a new ion composition model and a new spread F model (Bilitza and Reinisch, 2008). IRI-2012 improved the descriptions of the electron density, electron temperature and ion composition, (Bilitza et al., 2014). Finally, the latest version, IRI-2016, improved discrepancies between the F2 peak height and M(3000)F2 along with hmF2 improvements. Additionally there were general improvements for very low solar activity during the 2008/9 solar minimum (Bilitza et al., 2016).

IRI employs global spherical harmonic fits to the observations. These observations are spatially sparse, particularly in the southern hemisphere. These fits are performed at two R_{12} values, 0 and 100 while keeping foF2 constant above $R_{12} = 150$ (Bilitza, 1990, p. 52). It was inspired by the rapid ionospheric dynamics measured by early satellite research. Initially, four parameters were modelled, electron density, ion and electron temperature and relative positive ion densities (Bilitza, 2004). These ion densities were given as a relative amount to reduce inconsistencies between measurements. While there was a large amount of data for electron density there was a lesser amount for temperature, with only a small amount of ion data. These problems were further exacerbated due to large spatial and temporal data gaps in all forms of data resulting in a reduced accuracy in some regions. Additionally, most sites only produced the virtual height and didn't apply an inversion technique in order to obtain the true height. To gather height data two methods were initially available, incoherent scatter radar observations and topside sounder profiles. Unfortunately, there were only a few incoherent scatter radars available. These incoherent scatter radar techniques can be used to obtain measurement for the height of the foF2 layer, hmF2, but previously stated hmF2 measurements were severely lacking at the time. However, there was an abundance of M(3000)F2 observations available and since there exists a relationship between hmF2 and M(3000)F2 it was still possible to make global hmF2 maps. This relation was developed by Shimazaki (1955) and later refined by Bilitza et al. (1979) as shown in Equation 2.1. The two main methods of obtaining bottom side ionosphere data were ground based radio propa-

gation and rocket based measurements. IRI's results for this bottomside region was sufficiently different compared to data that a special symposium was called in order to discuss the issue before it was resolved. To obtain the plasma densities for the entire height range, multiple models were developed which were valid between various regions. Ion and electron temperatures were determined using a retarding analyser technique (Dumbs et al., 1979) along with incoherent scatter observations. From the data it was determined that the during the day electron temperature could be much higher than the ion temperature. Initially when determining electron temperatures they were found to be coupled to electron density but for the monthly median this relationship was negligible.

There have been numerous papers describing ionospheric monthly median processing but Jones and Gallet (1962) is especially relevant. This was one of the first papers to describe global ionospheric modelling using automated numerical techniques. Initially ionospheric maps were hand produced via an expert. However, this process is extremely slow, tedious and different experts would produce conflicting maps. Additionally, it required the use of unrealistic simplifying assumptions such as the longitudinal zonal system, which does not include the effect of longitude. Automatic analysis suffered from two main problems, the data was affected by noise, and the stations are irregularly spaced. These two issues still affect modern day mapping. Jones and Gallet (1962) describes a methodology for developing global maps. The diurnal variation is described using a Fourier analysis on the 24 hourly monthly median measurements and then each site was then corrected to the same solar local time. Then global variation of the Fourier coefficients was found using functions similar to Chebychev polynomials that were fitted to a two dimensional surface. This work resulted in global maps of various parameters for foF2, hmF2 and M3000.

Many papers have reviewed IRI (e.g. Bilitza et al., 2011, 2017; Pignalberi et al., 2016, 2018), and they agree that IRI produces accurate foF2 predictions. The latest version of IRI is IRI2016, which was extensively reviewed by (Bilitza et al., 2017). This paper discusses the latest improvements to IRI, including new models for hmF2 and improved topside ion density prediction capabilities. Additionally, it briefly discussed updates on modelling the real time ionosphere.

foF2 is fairly straightforward to measure with manual scaling with some complications for autoscaling. Critical height, on the other hand, is significantly more complicated as it requires scaling and inversion of the ionogram trace. However, it is possible to obtain hmF2 estimates from M(3000)F2. As M(3000)F2 is easier to measure than hmF2, it is often used to estimate hmF2. The equation (Bilitza, 1990, p. 53), governing this relation is,

$$hmF2 \propto \frac{1}{M(3000)F2 + foE/foF2} \tag{2.1}$$

where hmF2 is in km and the proportionality is dependent on solar activity functions related to the sunspot number and the magnetic dip latitude.

As Bilitza et al. (2017) discusses, there are three limitations with estimating hmF2 with the above approximation;

1. The limits in small scale modelling, due to the limited number of harmonics used.

2. The lack of available data.

3. The uncertainty in the M(3000)F2 to hmF2 relation from Equation 2.1.

The first limitation prevents IRI from reproducing the small scale features such as the sharp evening peak in hmF2 in low latitude regions. The third limitation causes inaccuracies due to the validity of the assumption, particularly during the recent solar minimum. As a result IRI 2016 added two new hmF2 models which do not rely on M(3000)F2. These are AMTB[1], an ionosonde based modelling effort, and Shubin-COSMIC which uses radio occultation data.

AMTB uses data from 26 sites which were part of the Global Ionospheric Radio Observatory (GIRO) network between 1998 to 2006 (Altadill et al., 2013). Global hmF2 values were obtained using spherical harmonics applied across modified geomagnetic dip angle and geographic longitude only at $R_{12} = 15$ and 120. Due to the lack of data, additional observations were added by using the relation that 15^o in longitude corresponds to one hour in SLT. This allowed the model to add 23 fictitious points across the globe per site. Spherical harmonics were then used to fit hmF2 across modified geomagnetic dip angle, μ, and longitude. Temporal estimates were then fitted using spherical harmonics. Overall this method has been reported to improve hmF2 RMS error on average by 25 % at low latitudes and 10 % at high latitudes when compared to the previous version of IRI2007s hmF2 model. The modified geomagnetic dip angle is defined as

$$tan(\mu) = \psi / cos^{1/2}(\Phi) \tag{2.2}$$

where ψ is the magnetic dip and Φ is latitude.

An alternative set of hmF2 coefficients, Shubin-COSMIC, was also implemented within IRI2016 (Shubin, 2015; Bilitza et al., 2017). These coefficients were generated from radio occulation data generated from three satellite's, CHAMP Satellite (2001-2008), GRACE (2007-2011) and COSMIC (2006-2012). Additionally 62 digisondes were used from the Digital Ionospheric Data Base of the Centre for Atmospheric Research of the University of Massachusetts, Lowell and Space Physics Interactive Data Resource (SPIDR). This ground based data spanned 1987 to 2012. This data

[1]Atadill Magdaleno Torta and Blanch

were filtered into two solar activity ranges $F_{107A} \leq 80$ and $F_{107A} \geq 120$, the remaining data $80 < F_{107A} < 120$ was used for validation. Where F_{107A} is the daily F10.7 index averaged over 81 days. The data was binned into according to modified dip (15 ° resolution) and geographic longitude (15 ° resolution). The modified geomagnetic dip angle was defined using an alternative definition as

$$\mu = arcsin\left(\frac{I_{300}(ryear)}{\sqrt{I_{300}^2(ryear) + cos(\phi)}}\right), \tag{2.3}$$

where ϕ is the geographic latitude, I_{300} is the geomagnetic field inclination at 300 km height and ryear is the decimal year:

$$ryear = year + (month - 0.5)/12. \tag{2.4}$$

The model was constructed similarly to AMTB with the mixture of spherical harmonics and Legendre polynomials. Overall this model produces accurate hmF2 values across the globe. There is a relative deviation less than 6 % when compared to data used within the model. In a comparison with independent data it is less than 8 % (Shubin, 2015).

In 2007 Zhang et al. (2007) reviewed IRI F2 peak parameters at Hainan Island (19.4 ^{o} N, 109.0 ^{o} E), a northern hemisphere low latitude site located near the peak of the northern equatorial anomaly. The monthly median foF2, hmF2 and M(3000)F2 observations used within this paper were obtained as hourly measurements from an ionosonde between March 2002 and February 2005. This time period within the descent of the solar cycle from the 2001 maximum to the 2008 minimum. This paper discovered that both CCIR and URSI foF2 values are systematically underestimated during the day between 5 % to 25 %, while IRI overestimated foF2 by up to 30 % pre-sunrise. Additionally, it performed better in 2002 than the other years. The comparison of hmF2 observations and IRI's M(3000)F2 model show poor agreement. The disagreement in CCIR coefficients was found to be due to IRI not producing the small scale structures. Figure 2.1 shows plots of observed hmF2 and CCIR predicted values for a variety of month and years. Within these plots there appears to be a lead in the model over observations that changes with season and year. The lag appears to be in the order of an hour, though it could vary more dramatically, but this lag was not discussed within the paper. The analysis in Chapter 4 also displays a lag and discusses it in detail.

IRI2012 foF2 variability was reviewed at Jicamarca during a variety of conditions (Adebesin et al., 2014). Jicamarca is a geomagnetic low latitude station at geographic coordinates (11.9 ^{o} S, 76.8 ^{o} W, 1 ^{o} dip) in South America. This study used automatically scaled hourly foF2 and hmF2 values with an error of $\pm$ 0.3 MHz at the 95 % confidence interval. The observations were from 2001 to 2002, 2004 to

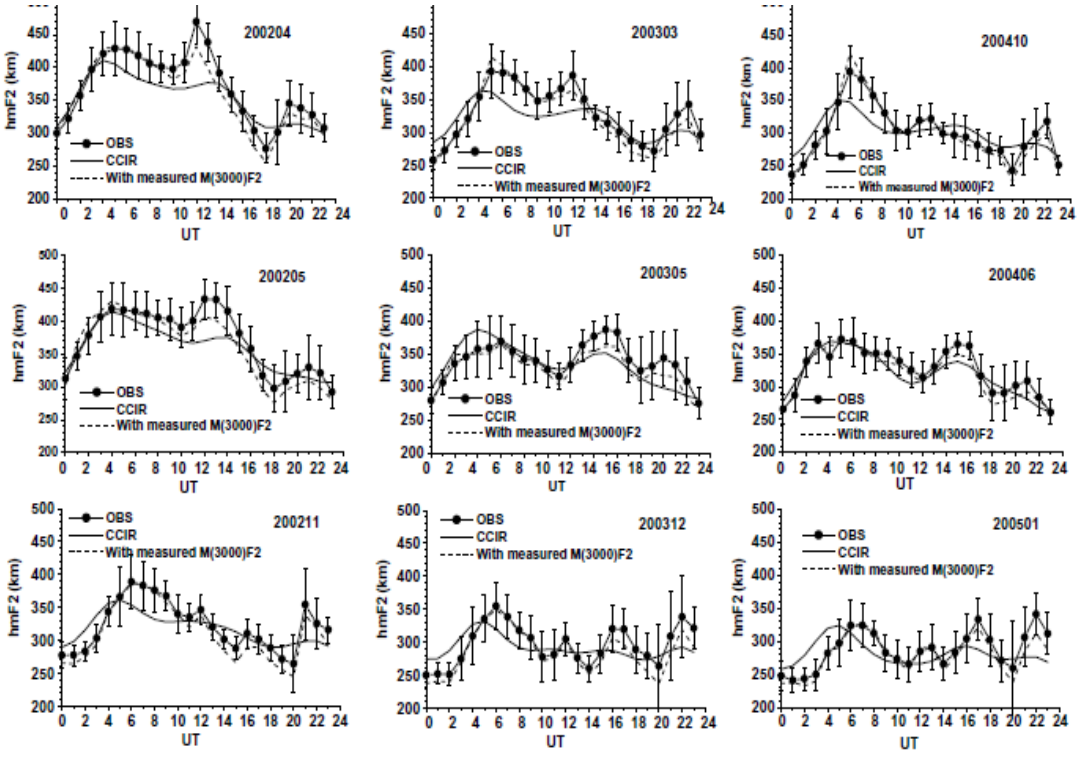

Figure 2.1: Figure 6 from Zhang et al. (2007) displays a comparison between IRI's CCIR hmF2 model, and IRI driven by M(3000)F2 and hmF2 observation. A lag in hmF2 predictions is particularly noticeable in December 2003 and January 2005.

2005 and 2009 to 2010 corresponding to high medium and low solar activity. Relative variability, V_R, was used within this analysis to compare IRI to observations. Relative variability is described using the monthly mean, μ, and standard deviation, σ, using

$$V_R(\%) = \left(\frac{\sigma}{\mu}\right) 100. \tag{2.5}$$

V_R is lowest during the day at 7 % to 16 % during the three regimes, and is higher at night between 20 % to 26 %, 14 % to 26 % and 10 % to 20 % during low, medium and high solar activity. Overall Jicamarca is well represented but an improvement is possible with the URSI options. Adebesin et al. (2014) shows comparisons between CCIR, URSI and observed values for a variety of seasons and years. Again, there appears to be a lag in IRI foF2 values with respect to observations which was not discussed within this paper.

IRI was also reviewed at Burkina Faso (12 ^{o} N, 1.8 ^{o} W) the low latitude African Ouagadougou ionosonde during solar cycle 22 between 1985 to 1995 (Ikubanni et al., 2014). The observations used within this analysis were obtained from a IPS-42 sounder. Ikubanni et al. (2014) generated a monthly mean model from monthly mean foF2 observations, which was compared to IRI. In Figure 2.3 a lag between model and observations could be seen when comparing both CCIR and URSI monthly median predictions to the monthly mean. The lag appears to be on the order of an hour.

Finally, a cross hemisphere analysis of IRI2012 was performed at two mid latitude sites Juliusruh, Germany and Hobart, Australia by Zhang et al. (2007). This analysis used manually scaled data from 1996 to 2009 measured at hourly intervals from the World Data Centre. This paper found that IRI performs well during the day time but tends to overestimate at night during high solar activity.

2.3 Summary

Ionospheric models generally operate at one of two distinct temporal levels: real time and climatological. Empirical models perform fits to historical observations, these models typically have difficulties due to the lack of available data, and that the data is typically sampled sporadically in time and space. Empirical models are useful as they are relatively simple to produce. Alternatively, physical principle models solve energy and momentum equations to calculate plasma densities etc. These models are extremely useful for examining the sensitivity of ionospheric parameters to a wide range of different conditions.

There are also two main temporal ionospheric models, real time and climatological models. Real time models try to simulate the current ionosphere (nowcasting) and immediate future (short term forecasting). This is useful for many applica-

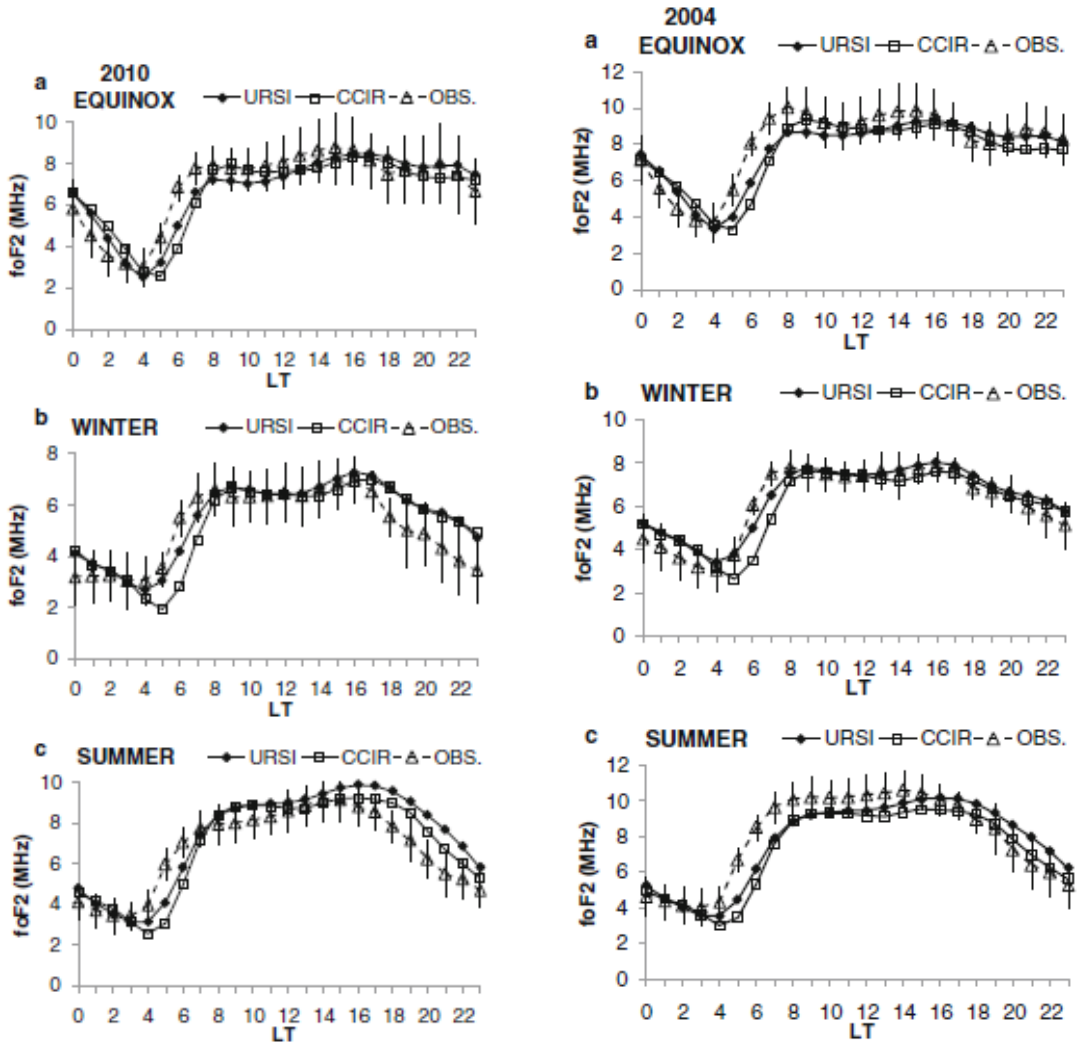

Figure 2.2: Figure 11 (left) and Figure 12 (right) from Adebesin et al. (2014) displays a comparison between IRI's URSI and CCIR foF2 model and observations. A lag in foF2 predictions is particularly noticeable in summer 2004 and equinox 2010.

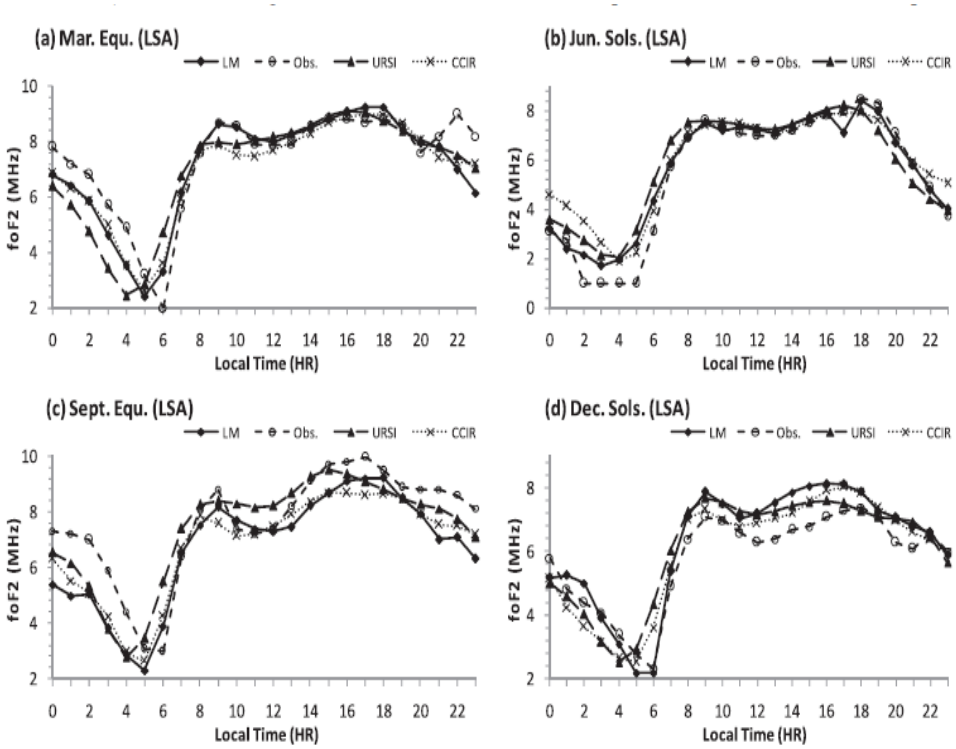

Figure 2.3: Figure 1 from Ikubanni et al. (2014) it provide a comparison between IRI's foF2 URSI, CCIR model, a local model's foF2 prediction and foF2 observations.

tions that require understanding of the current ionosphere. Climatological models predict the background ionospheric behaviour which is useful for examining large spatial structures over long time-scales. These models are useful for gaining a broad understanding of the ionosphere.

IRI is an global empirical climatological monthly median model of the ionosphere and is assessed in detail in Chapter 4. It has been continuously developed and reviewed since the late 1960s with numerous updates. For each hour and month IRI uses global spherical harmonics fitted across modified dip and longitude at two sunspot numbers, $R_{12} = 0$ and 100, keeping foF2 constant when R_{12} is above 150. Linear interpolation is used between these fits to provide a prediction for any sunspot number.

Multiple papers have reviewed IRI in the past in which they agreed IRI performs adequately. However, IRI tends to underestimate at solar minimum and overestimate at solar maximum. It has also been noted that IRI may have a one hour temporal lag at some sites. IRI's hmF2 model is typically less accurate than IRI's foF2 model.

The next chapter discusses the databases of foF2 and hmF2 used within this analysis. Additionally it describes the data filtering and sorting processes.

Chapter 3

Database Description

3.1 Introduction

The analysis in this thesis uses three scaled ionogram databases, viz. the World Data Centre (WDC), Global Ionospheric Radio Observatory (GIRO) and Jindalee Operational Radar Network (JORN). This chapter describes the temporal and geographic coverage provided by each dataset and examines the consistency between each datasets observations. Then it will detail how the observations were filtered to produce reliable monthly medians.

3.2 World Data Centre, Space Weather Services

The World Data Centre (WDC[1]) is a global consortium of data centres hosted by member institutes and contains a collection of scaled ionograms. Australian data was provided by the Space Weather Services (SWS) a section of the Australian Federal Government's Bureau Of Meteorology (BoM) in Sydney, Australia.

WDC has scaled foF2 observations from locations around the globe, with many countries having multiple sites. Unfortunately, the sounders are not uniformly spaced as seen in Figure 3.1, where sounder locations are in red and the geomagnetic dip equator and $\pm 30^\circ$ geomagnetic dip angles are in blue. The equatorial ionospheric anomalies, of interest in this study, usually peak near $\pm 30^\circ$ geomagnetic angles.

Some regions have a large number of sites e.g. Europe has an extremely dense network of sounders. However, there are significant gaps in the low latitude distribution particularly in oceanic areas, around Africa and South-East Asia. These regions are important as they cover the magnetic equatorial region, where the ionosphere is more dynamic than at mid latitudes. With data under-sampling near the dip equator, low latitude ionospheric modelling can be challenging.

[1] http://www.sws.bom.gov.au/World_Data_Centre

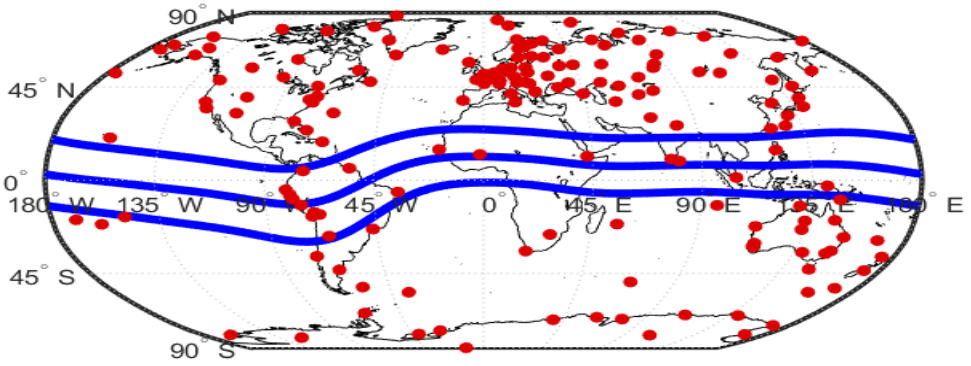

Figure 3.1: A map of WDC Sounders (red dots) with 0 ° geomagnetic dip equator and ± 30 ° geomagnetic dip angles (blue lines).

Figure 3.2 shows a map of Australasian region with the WDC available sounders in red. This region was initially the primary region of interest for this study but has since became a global review of IRI. While there is a large collection of sites on the Australian mainland, it is missing the important low latitude regions. The only low latitude sites are Manila (∼ 1 year of data), Vanimo (∼ 45 years of data) and Port Moresby (∼ 48 years of data). These last two sites have a large temporal extent, but there is insufficient spatial variation to accurately characterise this region.

The year range of WDC data depends on the site, with some sites collecting observations since the 1940s (e.g. Canberra). Other sites do not provide this much data: Darwin from 1982 to 2014, and Vanimo from 1984 to 2009. Fortunately the majority of sites contain one to two solar cycles. While many more sites exist globally, this analysis of IRI is primarily focused on these three sites due to the three solar cycles worth of data and their latitudinal span.

Australian WDC data has been manually scaled to URSI standards by the predecessor to SWS, the Ionospheric Prediction Service (IPS) (1947 to 2014) (IPS, 1994), and as a result it is expected to be reliable. If a problem or issue was detected during the manual scaling process a qualifier or descriptor was assigned to the data. These difficulties are denoted by different error codes and descriptors as listed by Piggott and Rawer (1978, p. 65), depending on the cause. For instance the error code D indicates the given observation is larger than the true observation, whereas N indicates conditions where a measurement cannot be interpreted. Other studies have used WDC data with the data filtered by removing all observations with an error code from the data (e.g. McNamara and Thompson, 2015). The filtering applied in

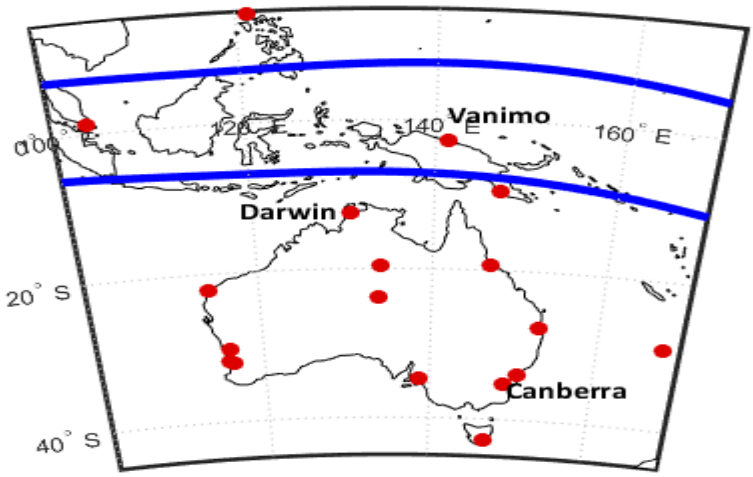

Figure 3.2: A map of WDC Australasian and nearby low latitude sounders (red) with 0 and -30 degree dip angles (blue lines), with latitudes and longitudes shown in geographic coordinates.

this study was less stringent as a majority of data is manually scaled and the results are expected to be accurate for many circumstances (e.g. spread F conditions). The filtering employed here only removed observations with the qualifier "I, M, T, U, A, D, E" or descriptor "A, C, D, E, M, N, P, S, T, V, W, Y, R". Scaled values were retained with the following qualifiers and descriptors, "J, O, Z", and "B, F, G, H, K, L, O, Q, X, Z". Tables, A.1 and A.2 in Appendix A, provide a complete description of the qualifiers and descriptors of the URSI scaling codes.

Figure 3.3 illustrates the effect of the filtering based upon certain qualifier or descriptors. The left plot shows the monthly mean amount of days of data available per hour at Canberra over the 34 year period, while the right plot shows the URSI parameter filtered observations. This process reduces the data from about 30 observations points per hour to 25 in some cases. The reduction in data availability is not uniform across the solar cycle as filtering appears to effect data availability at solar minimum more than at solar maximum years. Solar minimum years occur around 1986, 1996 and 2008 for solar cycles 22 to 24. During these times the filtering appears to remove several days of data. Additionally, there is a seasonal dependence, with more data at the solstices being removed than at the equinoxes. However, this effect is more subtle than the solar cycle effect and does not filter out large amounts of data.

The data were also filtered to remove days affected by of geomagnetic storms. The strength of geomagnetic storms are characterised by the Disturbance storm time (Dst) index, which is a measurement of the depression in the Earth's average horizontal magnetic field due to the storm-time magnetospheric ring current around

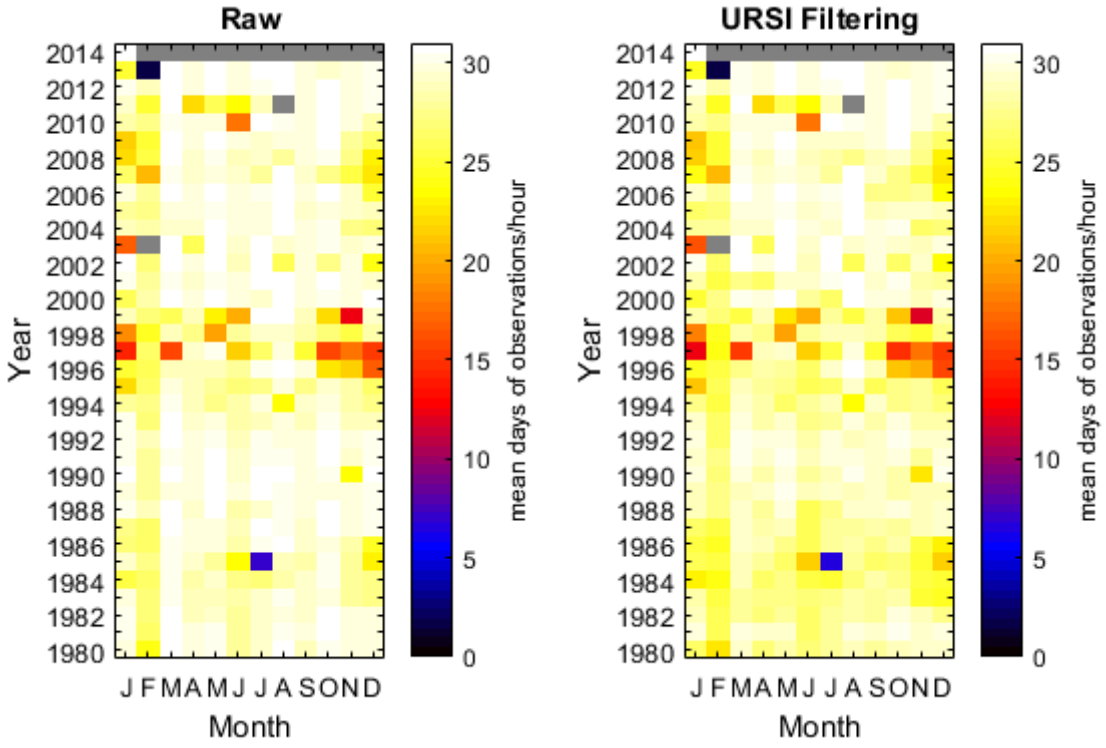

Figure 3.3: The monthly mean number of observations per hour of the unfiltered WDC data (left) and URSI parameter filtered WDC data (right) between 1980 to 2014, for Canberra. See text for details of the filtering process.

the equator, measured at near-equatorial geomagnetic observatories. Storm free periods typically have Dst in the range $\sim$ -20 nT to 20 nT. Moderate storms are defined as the maximum negative excursion of Dst of between -50 nT and -100 nT, strong storms between -100 nT and -250 nT, and super storms Dst is less than -250 nT. As we are interested only in a climatological model of the ionosphere we shall assess IRI with its storm model turned off and filter the data to remove the days when storms occurred. Typically other research papers filter storms by removing days where Dst is less than -50 nT. However, this analysis applied a more stringent filter to ensure that there was absolutely no influence from the recovery phase of a storm. This more stringent filter was chosen as the ionosphere usually takes a day to recover fully after the geomagnetic storm peaks. If the storm level for a particular day was Dst $\leq$ -75 nT then that day was removed from the analysis, and if Dst $<$ -100 nT the following day was removed as well. This ensures that only data for periods where the ionosphere has fully recovered are included in the analysis.

Figure 3.4 shows Canberra data availability with the additional Dst filtering. Clearly large amounts of data have been removed. Storms are less common during solar minimum than maximum, so less data is expected to be removed for years around 1986, 1997 and especially 2008, where solar cycle 24 had a deep solar minimum (Bilitza et al., 2012). This expected behaviour is seen in Figure 3.4. The Dst filtering removes more data than the URSI qualifier/descriptor filtering, which is also the case at Darwin and Vanimo (see Figures B.1 and B.2 in Appendix B). However, there is still enough data remaining after URSI flag and Dst filtering to calculate reliable monthly medians.

The minimum number of days required to produce reliable monthly medians was estimated by examining those months with almost an entire month of data after storm and URSI error code filtering. Monthly median foF2 was then calculated at each hour using 30, 20, 15, 10 and finally only 5 days of data. The monthly medians were then compared to determine the minimum number of days needed to produce a median in close agreement with that obtained using all the available data. This style of analysis was performed by Libo et al. (2003) at Wuhan and they found that a minimum of 13 observations were required for a reliable monthly median of hourly foF2 data.

Figure 3.5 shows the results of this analysis for Darwin foF2 observations at winter during solar maximum, July 2014. Clearly there is a good agreement between the medians for different sample sizes with only the 5 day median differing appreciably (up to 1 MHz) from the median derived using 30 days. This analysis was repeated for March 2008 (Autumn, solar minimum) at Darwin, with the results shown in Figure 3.5. Again there was a good agreement between the medians except for the 5 day median. These plots indicate that at least 15 days of data are required to ensure the calculated monthly median is reliable. Figure 3.6 displays the average

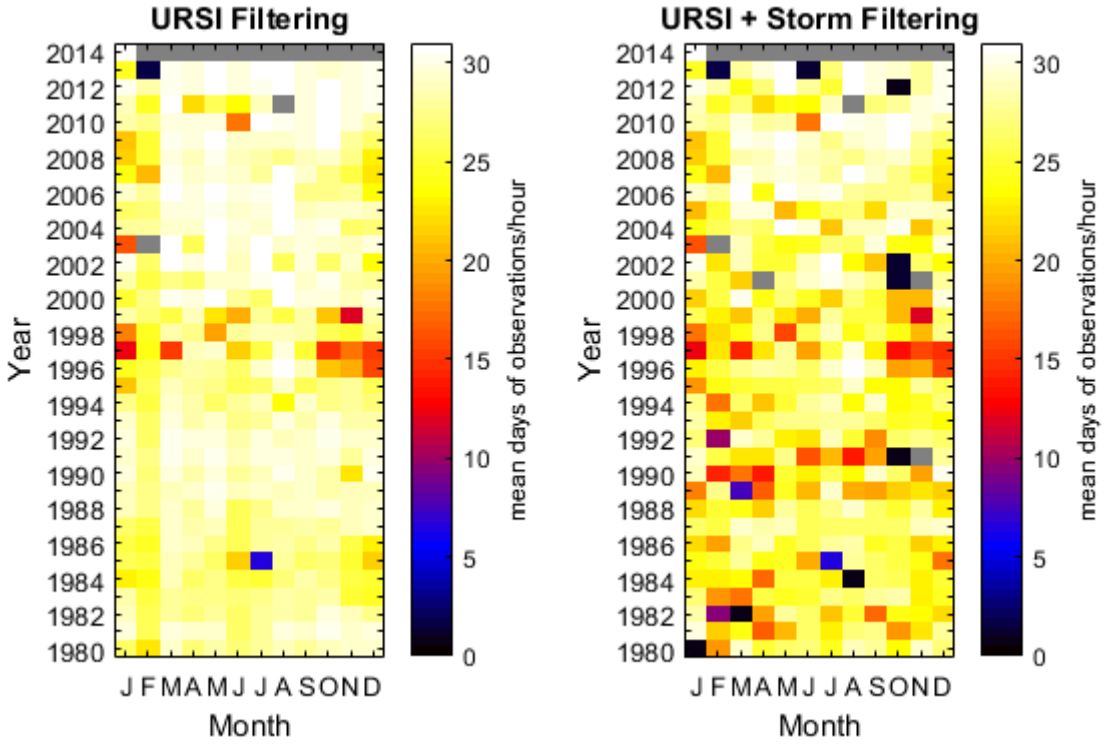

Figure 3.4: Average number of days contributing to the hourly calculations of the monthly median foF2 observations on the URSI parameter filtered WDC data (left) for Canberra and Dst plus URSI parameter filtered WDC data (right) between 1980 to 2014.

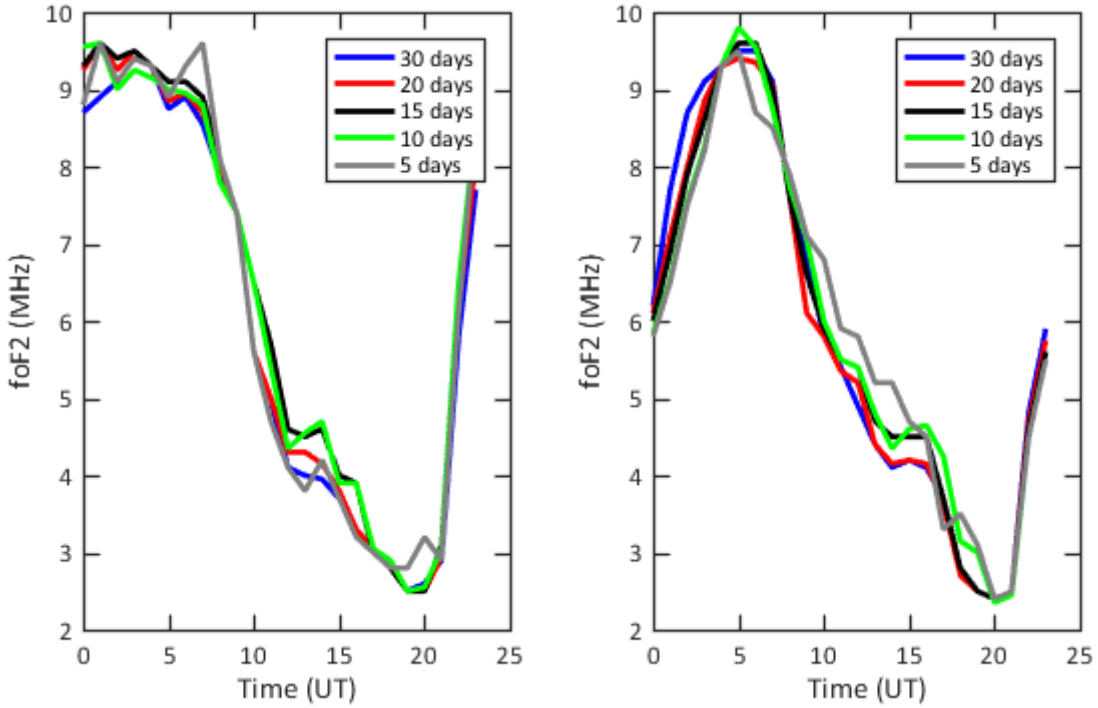

Figure 3.5: Monthly median foF2 for Darwin in July 2014 (left), and March 2008 (right) with a restricted number of days contributing (see text for details). Monthly medians were calculated after removing storms and URSI parameters.

number of days contributing for each hourly calculation of the monthly median for each month at Canberra, Darwin and Vanimo. We note that most months have greater than 15 days available data and so we expect that the monthly median foF2 in our subsequent analysis will be reliable.

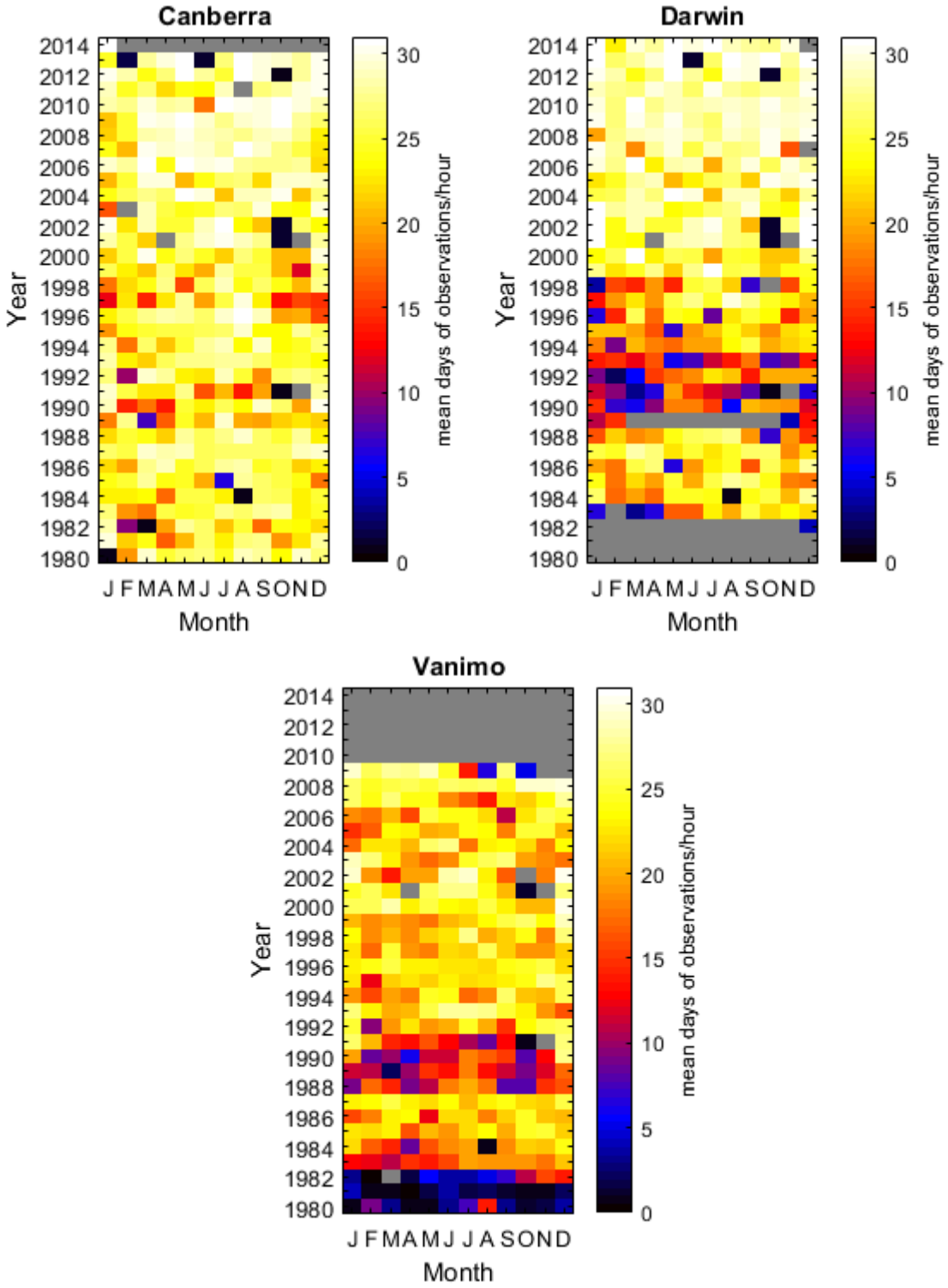

Figure 3.6: Average number of days contributing to the hourly calculations of the monthly median foF2 observations at Canberra (top left), Darwin (top right) and Vanimo (bottom).

Figure 3.7: A map of the GIRO sites as of 12/6/2018 available from http://ulcar.uml.edu/stationmap.html

3.3 Global Ionospheric Radio Observatory (GIRO)

The Global Ionospheric Radio Observatory (GIRO) is a global set of Digisonde ionospheric sounders (Reinisch and Galkin, 2011). GIRO is managed by the University of Mass Lowell, UML, U.S.A. with the help of various other institutions around the world. The GIRO database[2] provides a large amount of high resolution ionograms along with automatically scaled parameters. Figure 3.7 provides a map displaying current GIRO sites. Similar to the WDC map, there are a large amount of data in regions such as Europe, whereas others such as South-East Asia and central Africa lack sounders. Some other regions, such as around Japan, contain more sites than WDC. While GIRO provides a large amount of data it lacks sounders in the low latitude Australasian region.

GIRO provides high temporal resolution scaled data, with some sites producing ionograms at a 15 minutes cadence. In comparison, WDC produces ionograms at only hourly intervals. Unfortunately, the GIRO data only extends a few years back for most Australian sites. Canberra and Darwin data exists from 2016, while there are no Vanimo data at all.

Besides the lack of data from the Australasian sounders, the GIRO data also suffers from the use of automatic scaling. As described in Chapter 2, manual scaling is performed by an operator following guidelines specified in Piggott and Rawer (1978) and is expected to be robust. Automatic scaling of the GIRO soundings is performed

[2]This data can be accessed online at http://giro.uml.edu/didbase/scaled.php or using their SAO Explorer program which can be downloaded from http://ulcar.uml.edu/SAO-X/SAO-X.html.

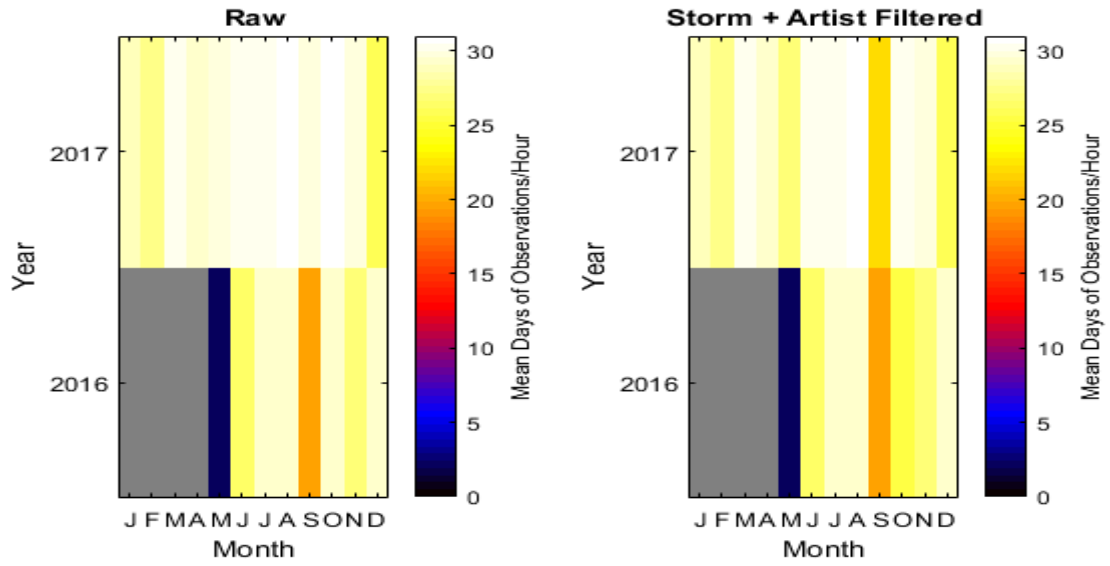

Figure 3.8: The mean number of daily observations in each hour of the monthly median. Describes the effect of the two filtering processes on GIRO data at Canberra. Left is the raw unfiltered data and right is the storm and confidence filter.

via the program ARTIST (Galkin et al., 2008), which follows the same guidelines. However, the automated algorithms are not expected to be as robust as a trained manual scalar. We note that a small amount of the GIRO data were manually scaled.

Each GIRO observation is saved with a confidence interval to represent the quality of the data. ARTIST saves the data with value between 0 to 100, where 100 represents high confidence in the data. Manually scaled data is saved with confidence score of 999. A confidence score of -1 is used to indicate that the method of scaling is unknown. After correspondence with Dr Ivan Galkin a minimum confidence interval of 50 was used to filter the data. The effect of this filtering level together with storm filtering can be seen in Figure 3.8. This figure shows the mean number of days of observations in each hour in the monthly median. Overall the reduction in to GIRO data availability due to filtering was minimal. Most months have enough data to calculate reliable monthly median foF2 as a function of hour.

As discussed in section 3.2, reliable monthly median estimation requires 15 days of data. When calculating the monthly median, for a given hour 15 different days of data was required.

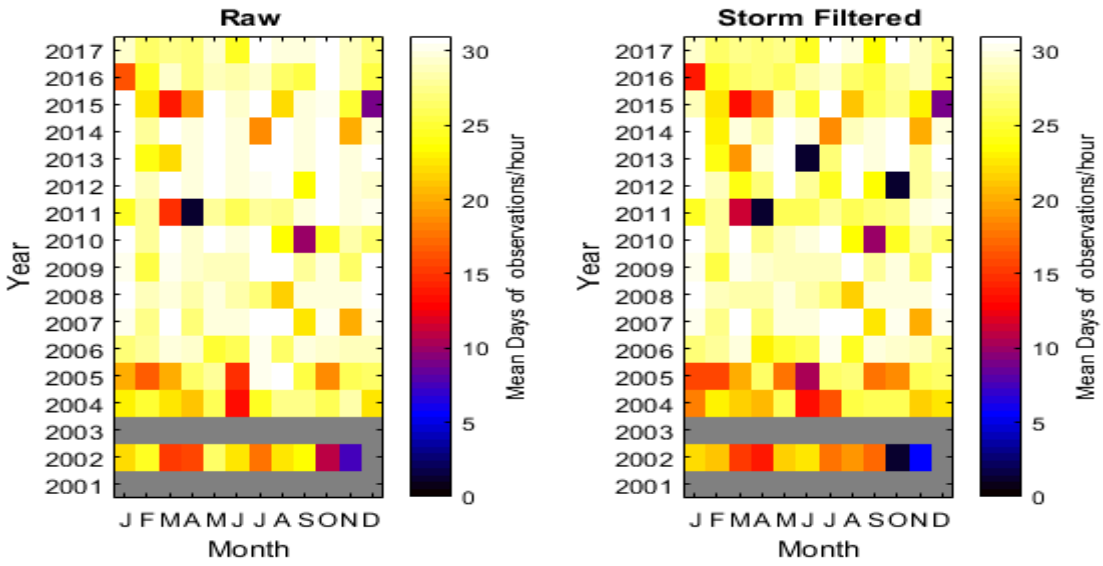

Figure 3.9: The mean number of days of observations per hour at Curtin. Left is raw data and right is storm filtered.

3.4 Jindalee Operational Radar Network (JORN) Ionosondes

The final database used in this analysis was the Jindalee Operational Radar Network (JORN) ionosonde data, supplied by Australia's Defence Science and Technology (DST) Group. DST Group is an Australian Federal Government entity in the Department of Defence. The High Frequency Radar (HFR) Branch within DST Group specialises in OTH radar and other HF technology research and development. The JORN sounder data used in the subsequent analysis covers 2002 through to 2017 at a cadence of up to 3.75 minutes from 13 sites distributed over Northern Australia.

Each of these sites in the JORN Sounder database have run over different years, and at different sampling rates at different times. These data were autoscaled without any scaling quality indicators. These data were filtered to remove storm time conditions as described previously. Figure 3.9 displays the effect of this filter at Curtin and shows that a most months have sufficient data to allow calculation of reliable monthly medians of foF2. As previously established, a 15 day minimum is required per hour for each monthly median.

3.5 Agreement between datasets

While each of the three datasets (WDC, GIRO and JORN) are useful they must be validated as a single data set to ensure consistency by comparing the datasets in the same months. However, there is limited overlap between the datasets: for instance, WDC and GIRO only overlap for a few years at Juliusruh, Germany, but utilize

the same sounder. WDC and JORN data overlap at Darwin, fortunately these are two independent sounders. Unfortunately, there is no overlap between GIRO and JORN. The limited over overlap between the datasets makes investigation of the agreement between them difficult.

Initially, WDC and GIRO were compared at Juliusruh between 2001 to 2014. At Juliusruh there is only one sounder, so this comparison is mainly useful for comparing the different scaling techniques (GIRO observations are automatically scaled while WDC were manually scaled). These comparisons are displayed in Figure 3.10, and it can be seen there is a excellent agreement between the two datasets. They differ by only a few hundred kHz at most and usually by far less, indicating that the two scaling methods produce comparable results. The root mean square of the difference between the two datasets from all available Juliusruh data is 0.2 MHz, 0.2 MHz and 0.3 MHz for the monthly median, tenth and ninetieth percentiles respectively. Overall, WDC and GIRO observations agree well.

A comparison between WDC and JORN data at Darwin is displayed in Figure 3.11 and there is a reasonable agreement between the two datasets. However, there are periods of large disagreements (up to approximately 1 MHz). This may be to the scaling methods implemented for the JORN sounder data. The RMS difference between the two datasets is 0.4 MHz, 0.5 MHz and 0.5 MHz for the monthly median, tenth and ninetieth percentiles respectively. These two datasets have a larger disagreement than WDC and GIRO but overall, there is still good agreement between WDC an JORN. Unfortunately, there is no overlap between JORN and GIRO datasets and so comparisons between these two datasets cannot be made.

It is expected that WDC is the most accurate source, particularly in the Australian region, as it is manually scaled. GIRO and JORN data is expected to be less accurate due to the use of automated scaling. WDC data was used as first preference then GIRO, then JORN. This should provide the most accurate medians of the ionospheric parameters.

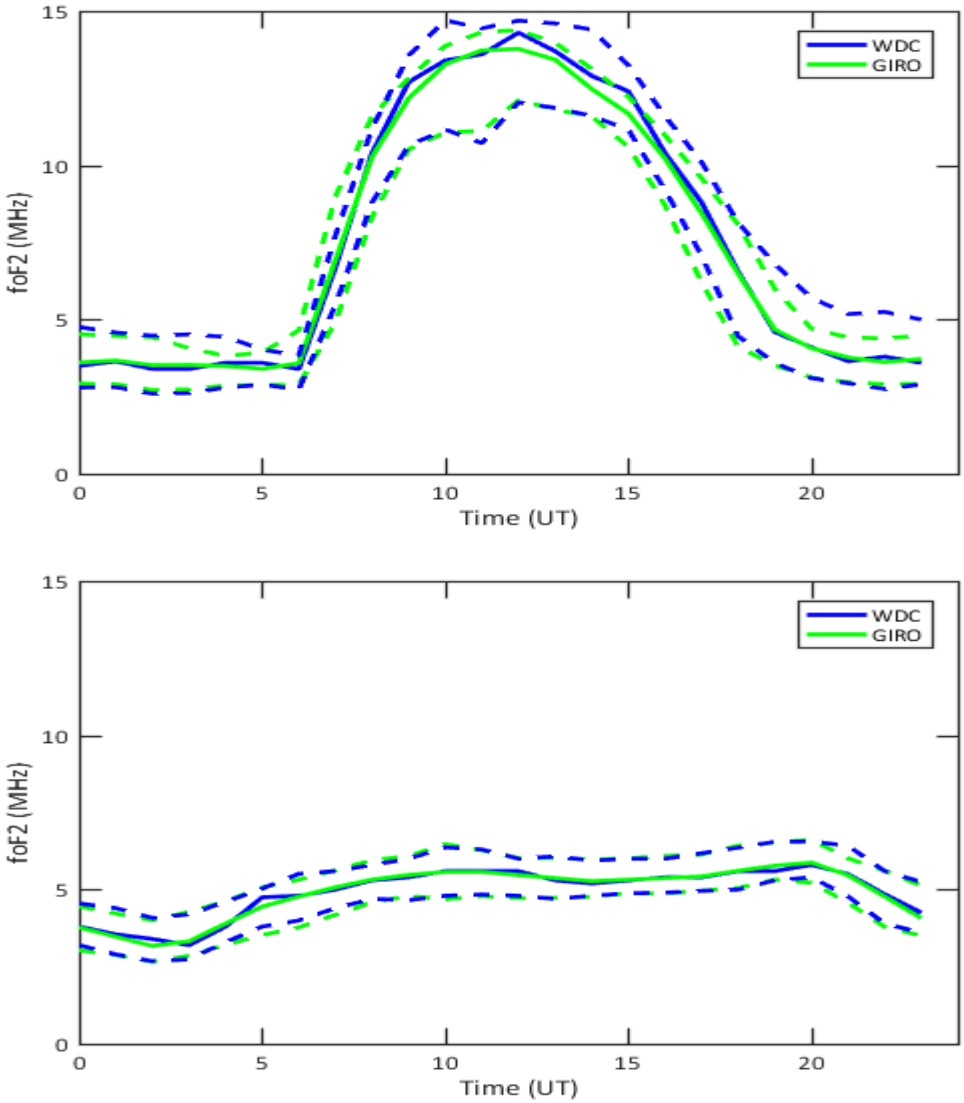

Figure 3.10: A comparison of GIRO (green) and WDC (blue) monthly medians and the tenth and ninetieth percentiles at Juliusruh. Top is January 2002, bottom is May 2006.

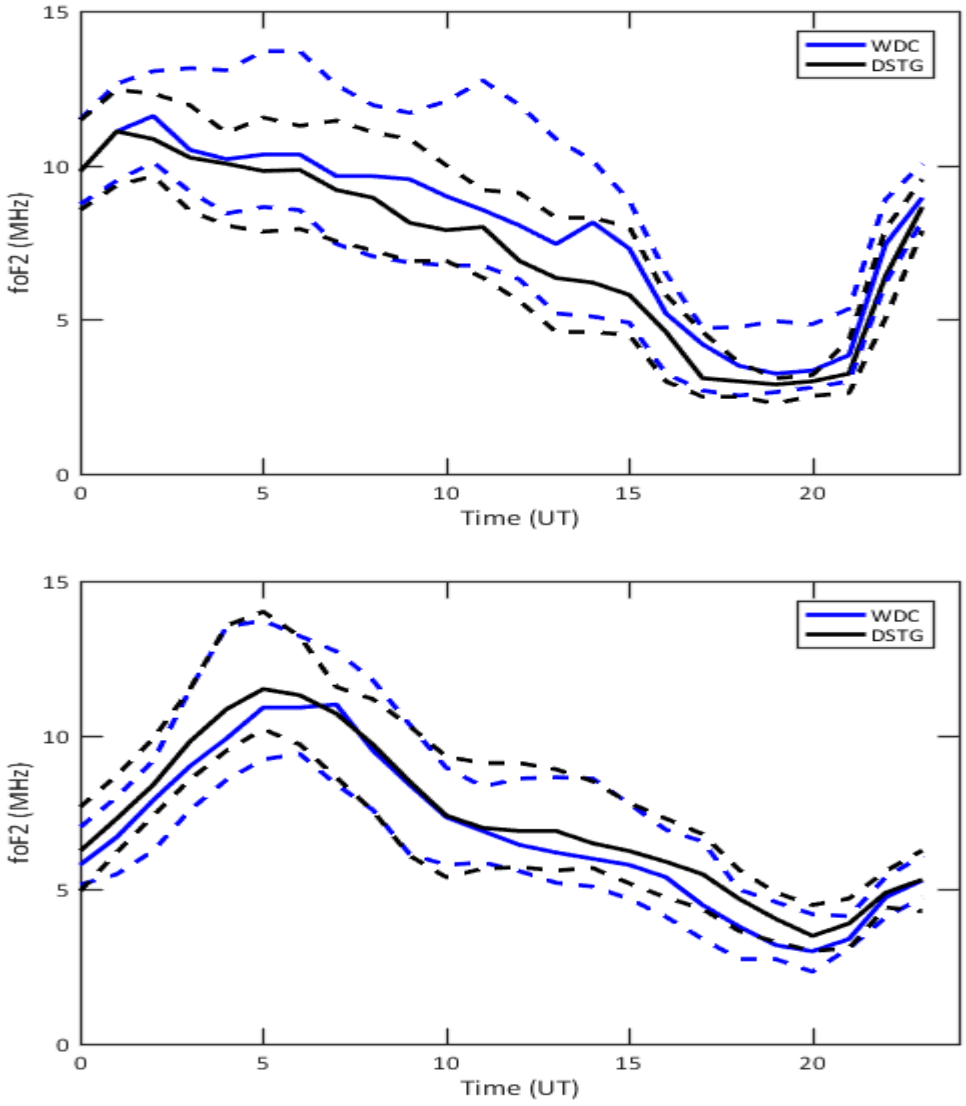

Figure 3.11: A comparison of JORN (black), and WDC (blue) monthly medians and the tenth and ninetieth percentiles at Darwin. Top is September 2003, and bottom is January 2007.

3.6 Combined Data

Combining all three datasets improves the spatial and temporal resolution and results in 188 sounders across the globe, albeit with non-uniform distribution. The sites tend to be mostly located in the northern hemisphere, primarily in Europe and the Americas with Australia as the main southern hemisphere contributor. Many other regions lack sufficient sounders for detailed coverage. The combined dataset still lacks sufficient measurements from the low latitude regions to fully account for the steep foF2 gradients there, but it is significantly better than any solo dataset. Figure 3.12 displays all the sites used in the combined data set.

The three datasets do not fully overlap in time. WDC tends to contain older data with not much available past 2014, while GIRO tends to contain newer data (from 2016 onwards) and the JORN sounder database contains data from 2002. Additionally, data availability tends to be reduced during solar maxima due to the effect of storm removal. Observations were predominantly from the last 20 years covering solar cycles 22, 23 and 24. However, some sites have data from 1940. Despite these limitations this dataset should provide an accurate and useful description of the background ionosphere to adequately test IRI or develop a new model.

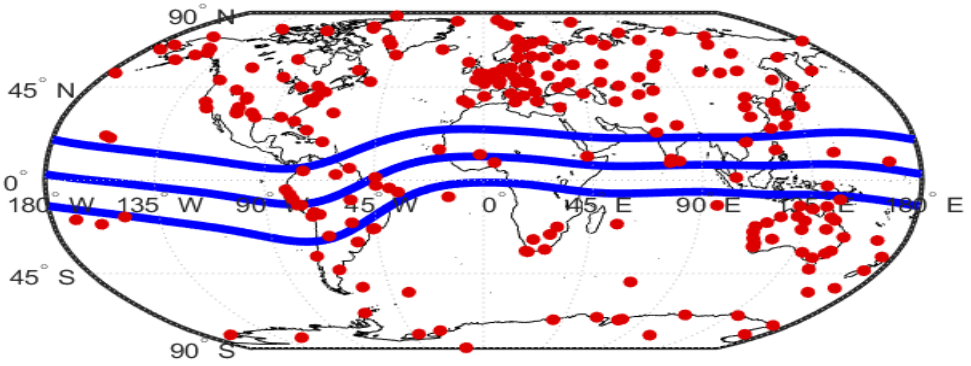

Figure 3.12: A world map containing the combined WDC, GIRO and JORN dataset. Each sounder is in red, with the 0 and ± 30 ° geomagnetic dip angles in blue.

Year	S_{12}	N Observations
1963	28.7	22
1964	22.6	27
1965	25.4	31

Table 3.1: The number of foF2 observation at Adak in January in the S_{12} bin = [20, 40] at 2 UT.The mean S_{12} for this hour is 25.4.

3.7 Data Interpolation

To aid the interpretation of the storm and URSI qualifier filtered daily observations, the data was distributed into five sunspot bins, S_{12} = [0, 20], [20, 40], [40, 80], [80, 150], [150, 250]. All the available data from 1940s onwards were binned in this manner. The monthly median, tenth and ninetieth percentiles were then calculated for each hour. For example, at Adak, Alaska during January the S_{12} = [20, 40] bin at 2 UT contains data from the years 1963, 1964 and 1965. The S_{12} values for these 3 months are 28.7, 22.5 and 25.4 respectively. There is a total of 80 observations in these three years. The mean S_{12} for these three years during January is 25.4. This is summarised in Table 3.1.

It is desirable to interpolate the calculated monthly median foF2 data to the centre of the respective S_{12} bins. This was done using linear interpolation, as displayed in Figure 3.13. Other authors have discussed that a 2nd order fit may be better (Mahajan et al., 1997), however for our analysis linear interpolation was chosen due to simplicity and the linear behaviour displayed in Figure 3.13. This fit was repeated for all hours, months and each S_{12} bin, to produce the interpolated observations at

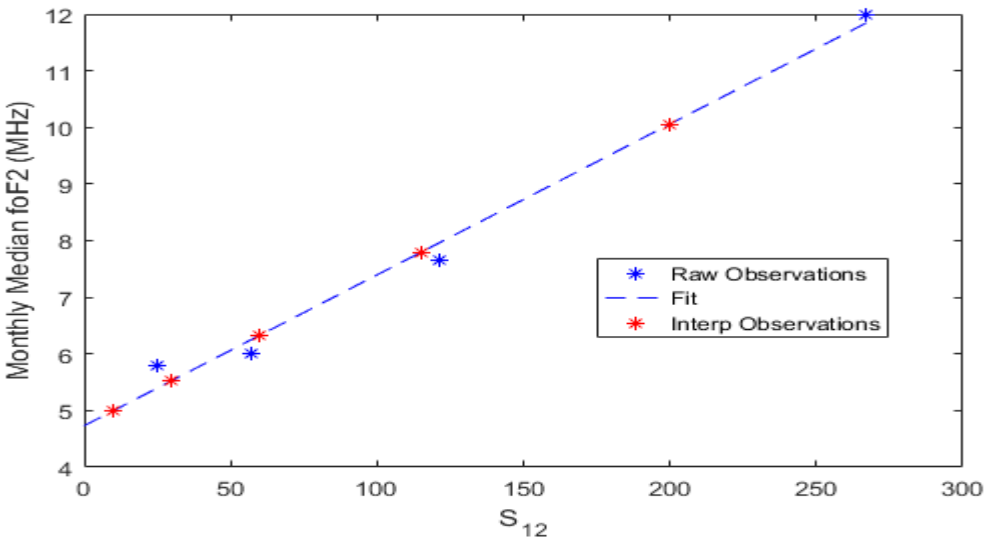

Figure 3.13: An example of the linear interpolation of the raw observations (blue) to produce the interpolated monthly median foF2 observations (red) for Adak, Alaska in January at 2 UT.

$S_{12} = 10$, 30, 60, 115 and 200. The final results for Canberra is shown in Figure 3.14.

The filtering process was repeated on hmF2 observations from the GIRO and JORN databases. WDC observations of hmF2 were not included as this database does not provide these observations.

3.8 Summary

This analysis used a collection of sounder observations from the World Data Centre, Global Ionospheric Radio Observatory and Jindalee Operational Radar Network. These datasets were compared to determine the agreement between the sets. It was found that the datasets agree fairly well. A combined dataset was produced by preferentially adding WDC, then GIRO and then JORN.

These data were filtered to remove the effects of poor scaling and ionospheric storms. To remove poorly scaled observations WDC was filtered by removing a selection of qualifiers and descriptors and GIRO was filtered using a minimum confidence level. Storms were filtered by removing and day of observations with Dst $\leq$ -75 nT, if the day had a Dst $\leq$ -100 nT the next day was also removed. The monthly medians were then calculated with the requirement of at least 15 observations from different days per hour.

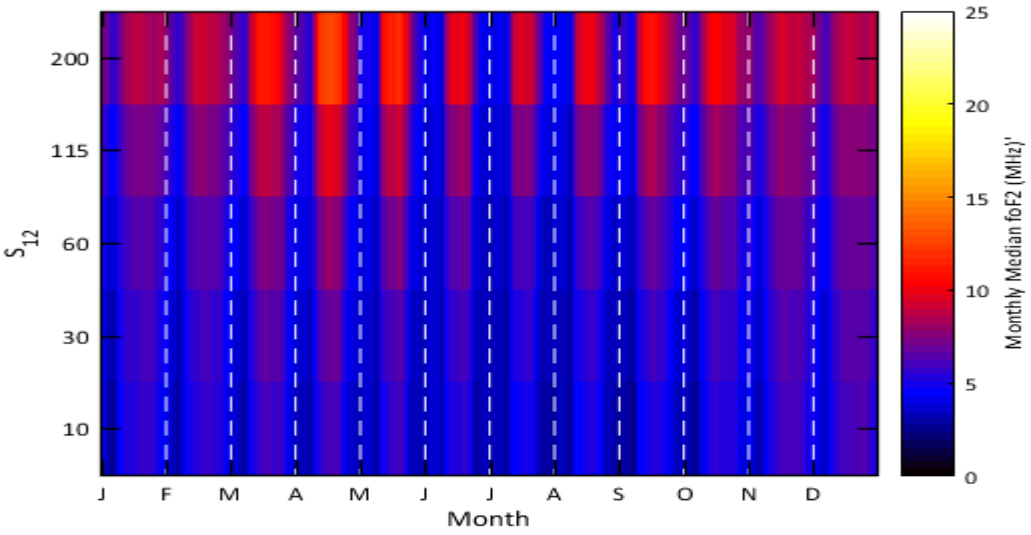

Figure 3.14: Interpolated median foF2 observations at Canberra as a function of S_{12}, month and time of day. The horizontal axis is 24 hours nested within month. White dashed lines indicate different months.

To aid the interpretation the data were binned into sunspot bins, $S_{12} = [0, 20]$, $[20, 40]$, $[40, 80]$, $[80, 150]$, $[150, 250]$. Linear interpolation was used to interpolate the data to the centre of each bin.

Chapter 4

IRI Analysis

IRI is a climatological model of the ionosphere that is capable of producing global monthly median predictions of various ionospheric parameters. However, the accuracy of this model must be validated using global observations.

Before we develop the new model of foF2 and hmF2 described in Chapter 5, it is necessary to assess the existing climatological model IRI to firstly determine if there are any issues with it which can be improved. The analysis methodology in this chapter is used to assess the new model and compare it with IRI. IRI must be tested in a variety of conditions, such as different parts of the solar cycle, seasons and time of day. It should be tested geographically by examination of a variety of latitudes and longitudes to measure the global structure of the errors.

4.1 Residual Analysis

4.1.1 Canberra

IRI was initially tested at a mid latitude site, Canberra, Australia (Lat = -35 ° N, Long = 149 ° E, Dip = -66 °). Firstly, a few specific months were examined by comparing IRI's URSI foF2 monthly median prediction with the monthly median foF2 WDC observations as shown in Figure 4.1. This figure shows IRI's performance in January 2001, which is solar maximum conditions. Clearly, IRI performs well in this case as it captures the ionospheric variability with only small deviations. However, it appears to over predict after 20 UT by a maximum of approximately 1 MHz. This overestimation occurs immediately post sunrise.

This analysis was repeated during the winter of a solar minimum year, June 2009, and is shown in Figure 4.2. Sunset occurs at 7 UT and sunrise occurs at 21 UT. Again IRI follows the general tends of the foF2 closely during the day. However, it underestimates the night time foF2 by a considerable margin.

Analysing single months can make it difficult to identify large scale trends. To determine the diurnal, annual and solar cycle dependence of any IRI errors, the residuals between IRI derived ionospheric parameters and the monthly median ob-

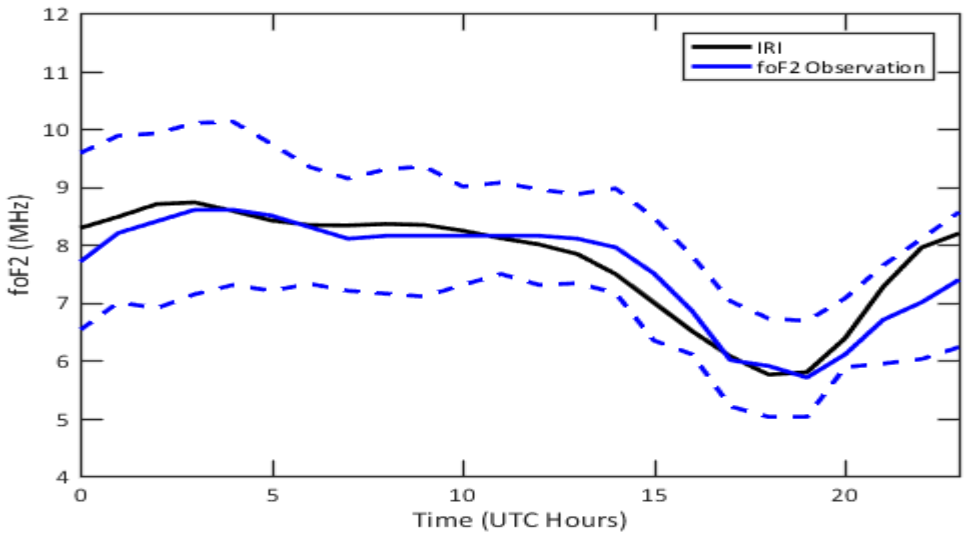

Figure 4.1: A comparison between IRI (black line) and Canberra medians (blue line) and tenth and ninetieth percentile (dashed blue lines) from WDC observations for January 2001. Sunset occurs at 9 UT and sunrise occurs at 19 UT.

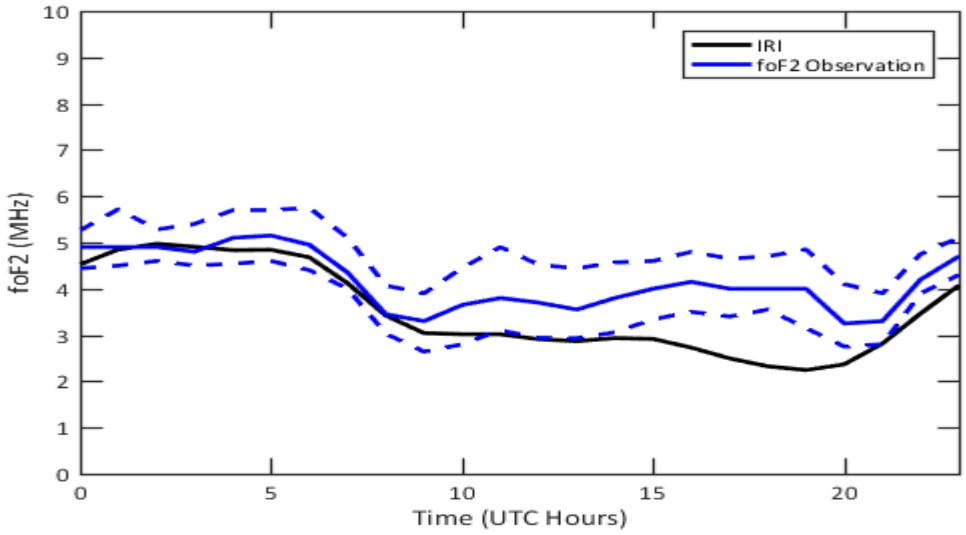

Figure 4.2: A comparison between IRI (black line) and Canberra medians (blue line) and the tenth and ninetieth percentile (dashed blue lines) from WDC foF2 observations for June 2009. Sunset occurs at 7 UT and sunrise occurs at 21 UT.

servation of that parameter were calculated at each hour of the day. The residual is defined in Equation 4.1:

$$\text{Residual} = 100\left(\frac{\text{IRI} - \text{Observation}}{\text{Observation}}\right)(\%). \tag{4.1}$$

This residual analysis was performed at Canberra for the foF2 data over three solar cycles and is shown in Figure 4.3. The horizontal axis of this figure contains 24 $\times$ 12 residuals which represents each hour in the monthly median for each month.

From Figure 4.3 it is clear there is a diurnal, seasonal and solar cycle structure in the residuals. The residual for each hour is not uniform throughout the month and will underestimate or overestimate at different times of day, as seen clearly in winter, 1993. IRI tends to underestimate in winter (June to August) and overestimate in summer (December to February). During solar minimum years it tends to consistently underestimate foF2 while at solar maximum IRI overestimates.

The mean and standard deviation of the foF2 residual was calculated for each month in each year and displayed in Figure 4.4. The standard deviation tends to be larger during solar minimum years. The standard deviation is also larger in winter compared to summer. During solar minimum IRI tends to underestimate foF2 as shown by the negative mean, and slightly overestimates foF2 in summer.

Figure 4.4 displays a clear seasonal and solar cycle relationship in both the mean and standard deviation of the foF2 residuals. However, it is still difficult to examine the solar cycle effect directly. This was solved by binning the data into S_{12} bins to produce the interpolated medians as described Chapter 3 (see Figures 3.13 and 3.14). The residuals between IRI and these interpolated medians is illustrated in Figure 4.5. This figure clearly displays the solar cycle dependence of the foF2 residuals. It indicates that during solar minimum years IRI tends to underestimate foF2 while during solar maximum years IRI tends to over estimate. During solar minimum it appears that IRI performs slightly better at predicting foF2 during summer than winter. This behaviour agrees with the trends displayed in Figure 4.3.

To further aid the interpretation of the seasonal and solar cycle behaviour of the IRI residuals, the mean and standard deviation of the residuals are calculated for each month and displayed in Figure 4.6. This figure clearly displays the seasonal and solar cycle trends of the IRI residuals. However, this is at the expense of losing information regarding the diurnal behaviour of the residuals. This is discussed later in Section 4.3. It is important to note that these plots should be examined together. A low mean residual could be due to either IRI agreeing well with the observations or that negative residuals at particular times of the day are cancelled out by positive residuals at other times of the day. If the former, then the standard deviation of

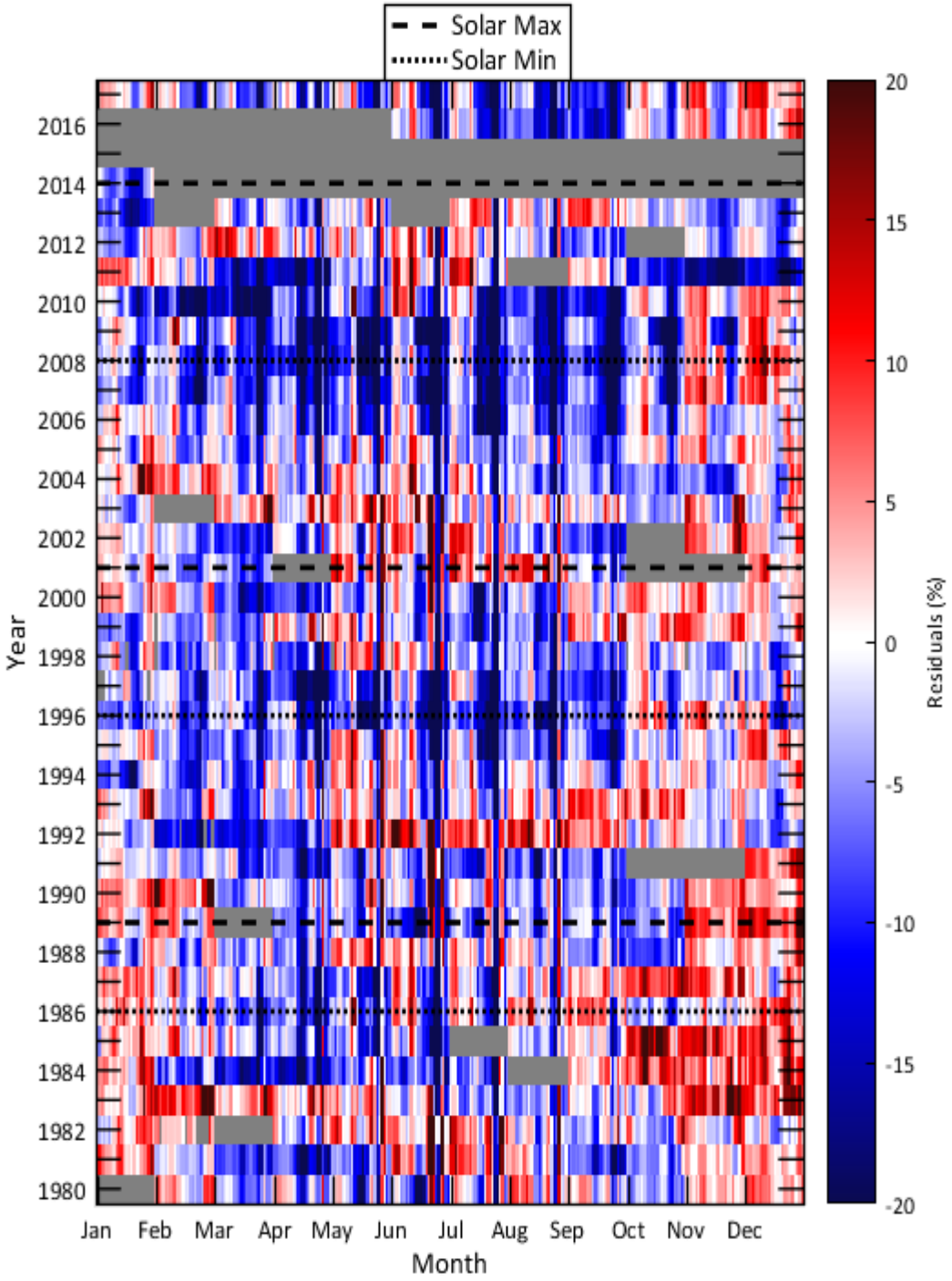

Figure 4.3: The foF2 residuals between IRI and the monthly median observations calculated for each hour at Canberra. The hourly residuals are nested within month. Also indicated are the solar maximum years (long dash line) and solar minimum years (short dash line).

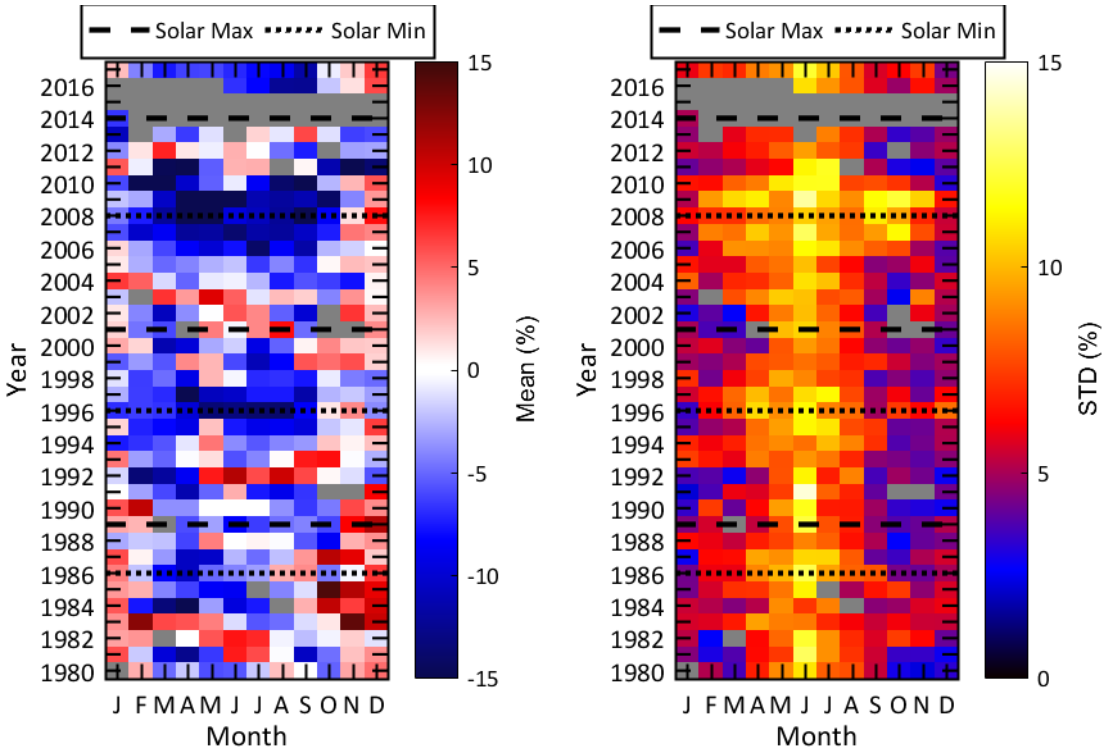

Figure 4.4: The mean residual (left) and standard deviation (right) between IRI and median foF2 observations at Canberra. Also indicated are the solar maximum years (long dash line) and solar minimum years (short dash line).

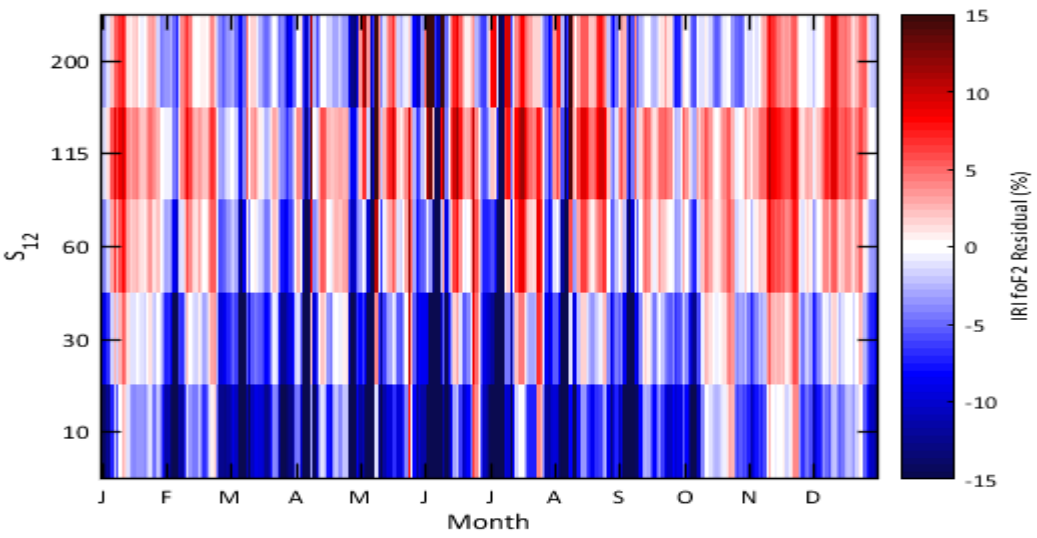

Figure 4.5: Residuals between IRI and foF2 observations for each hour in the month at Canberra.

the residuals will also be low. If the latter then the standard deviation will be large. The April S_{12}=115 bin in Figure 4.6 illustrates this; the mean residual is low (0.2%) whereas the standard deviation is large which indicates that for this epoch while there is no bias in the IRI foF2, it does not characterise the diurnal behaviour as well as one might like.

It is desirable to reduce these data further to produce a simple set of metrics to describe the performance of IRI at each site. This allows the performance of IRI at various locations to be easily compared. The root-mean-square (RMS) of the mean and standard deviation of the residuals was chosen as the metrics. At Canberra these values are 5.0 % and 6.0 % respectively.

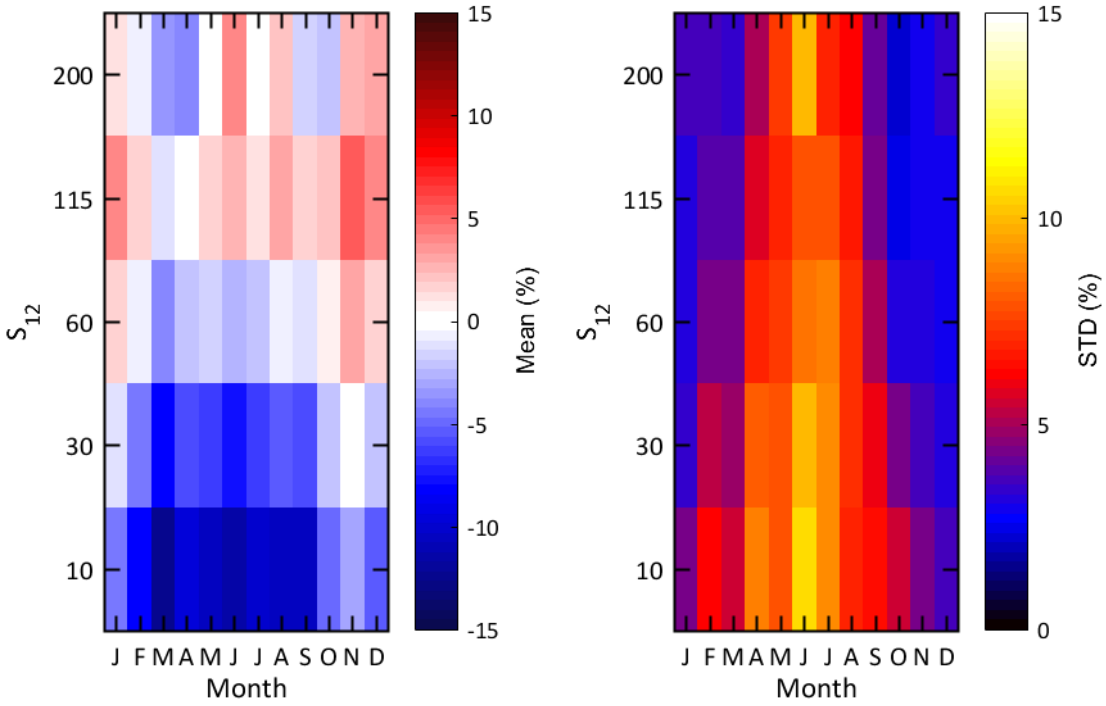

Figure 4.6: The mean (left) and standard deviation (right) of the IRI foF2 residuals as a function of S_{12} and season at Canberra.

4.1.2 Darwin

The analysis described above was repeated at Darwin (Lat = -12.4 ° N, Long = 130.9 ° E, Dip = -40 °) a low latitude site on the southern edge of the southern equatorial anomaly. Figure 4.7 displays the mean and standard deviation for the IRI foF2 residual error for comparison with Figure 4.6. IRI clearly performs worse at Darwin than at Canberra, which may be expected as the low latitude ionosphere has greater complexity, due to the equatorial anomaly, than at mid latitudes. Examination of the mean residuals clearly shows that IRI tends to overestimate foF2 at Darwin during winter at high sunspot numbers and underestimate in summer at low sunspot numbers. The RMS mean and RMS standard deviation in the foF2 residuals are 8.9 % and 10.3 % respectively. In comparison, Canberra has a RMS mean foF2 residual of 5.0 % and a RMS standard deviation of 6.0 %.

4.1.3 Vanimo

IRI's foF2 model was also tested at Vanimo (Lat = -3 ° N, Long = 141 ° E, Dip = -22 °) a low latitude site situated under the southern equatorial anomaly peak. Results are summarised in Figure 4.8. The RMS mean foF2 residual and RMS standard deviation for Vanimo are 4.9 % and 9.7 % respectively. Overall, IRI's foF2

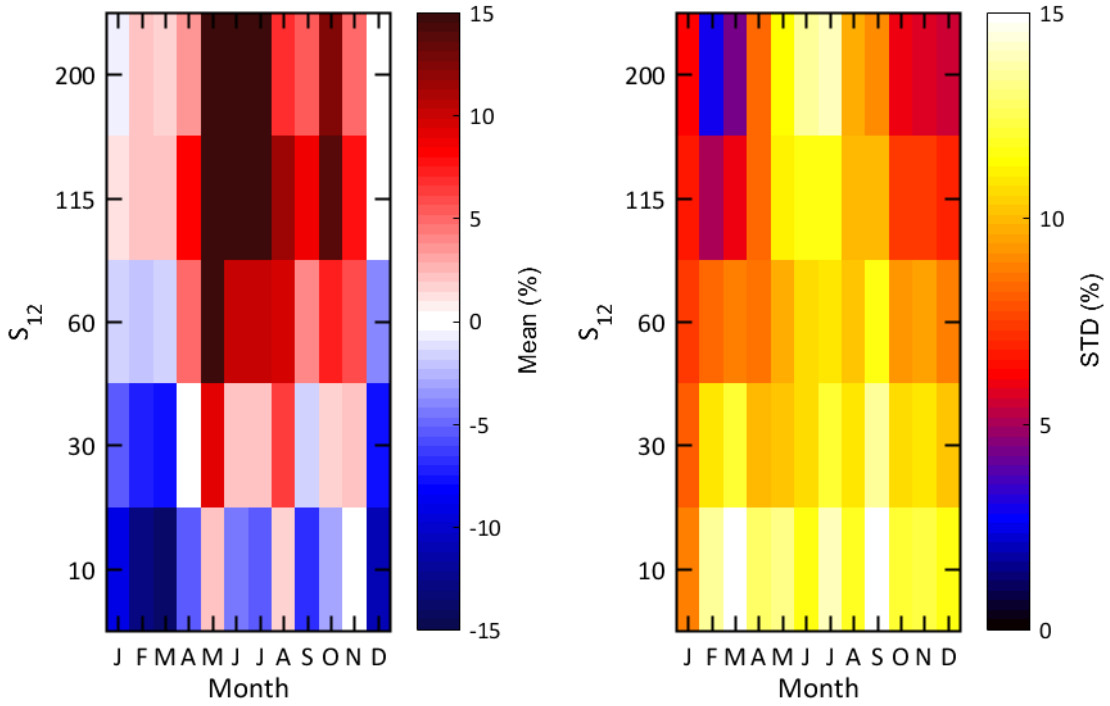

Figure 4.7: The mean (left) and standard deviation (right) of the IRI foF2 residuals as a function of S_{12} and season at Darwin.

model performs better at Vanimo than at Darwin but not as well as Canberra.

The mean foF2 residual in Figure 4.8 at Vanimo show a seasonal and solar cycle influence. IRI underestimates foF2 from December to August for low sunspot numbers (S_{12} = 10, 30). IRI overestimates during May to December for mid sunspot numbers (S_{12} = 60, 115).

The standard deviation of IRI's foF2 residuals displayed in Figure 4.8 also displays a clear solar cycle influence. The standard deviation is highest at low sunspot numbers and gets progressively lower at high sunspot numbers. There may be a weak seasonal influence on the standard deviations at solar minimum, which correlates with Canberra results.

Site	Dip (°)	RMS Mean (%)	RMS STD (%)
Canberra	-66	5.0	6.0
Darwin	-40	8.9	10.3
Port Moresby	-33	9.2	9.7
Vanimo	-22	4.9	9.7
Jicamarca	1	6.4	10.2
Djibouti	11	5.6	10.6
Puerto Rico	46	12.3	8.0
Tokyo	49	7.7	7.8
Juliusruh	69	6.5	5.6

Table 4.1: Table summarising the RMS mean and standard deviations of the foF2 residuals between IRI and observations over the entire solar cycle for a selection of sites ordered by geomagnetic dip angle.

Table 4.1 summarises the results of IRI's foF2 performance at a variety of sites across the globe, by reducing the site variation into two metrics: the RMS of the mean foF2 residuals for each month and RMS of the standard deviation. Overall, IRI appears to perform better at mid latitudes than in low latitudes. It appears the foF2 residuals are largest at the northern edge of the northern equatorial anomaly, and in the southern edge of the southern equatorial anomaly (i.e. the 'outside edges' of the anomalies) with values of 12.3 % and 9 % respectively suggesting greater variability there than at other locations.

4.2 Global Analysis

IRI's residual errors are explored further in global plots of the ionosphere using a large selection of sites. The data from each site have been temporally shifted to the same solar time. This was done using the solar zenith angle, the angle between the centre of the Sun's disk and zenith (Davies, 1990, p. 65) as, to first order, the F2 region electron density (and hence foF2) is solar controlled. First it was determined when the Sun was at the highest point in the sky. Then a circular shift was performed on the data to set this peak at noon solar local time (LT). Figure 4.9 displays an implementation of this procedure for March at low solar activity ($S_{12} = 10$). In this figure foF2 from all sites is plotted as a function of modified dip and it clearly shows the latitudinal structure of the ionosphere, the two equatorial anomaly peaks can be seen at $\sim \pm 25\,^\circ$. The mid latitudes can also be seen between -60 ° and -40 ° (40 ° and 60 °) for the Southern (Northern) Hemisphere.

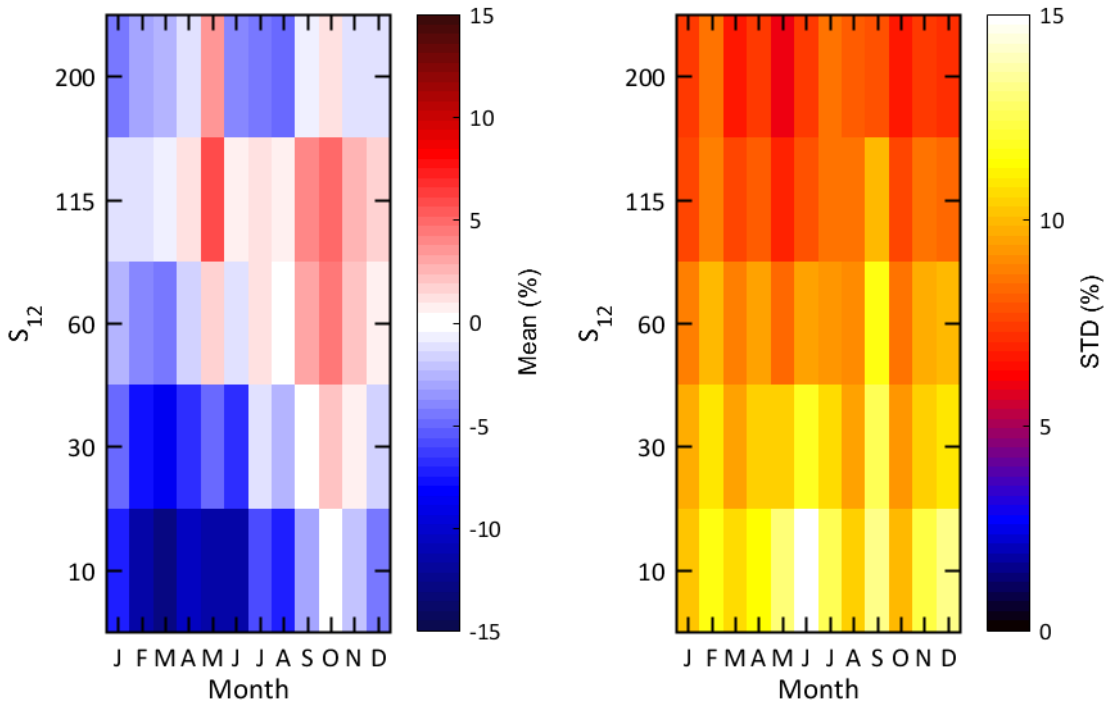

Figure 4.8: The mean (left) and standard deviation (right) of the IRI foF2 residuals as a function of S_{12} and season at Vanimo.

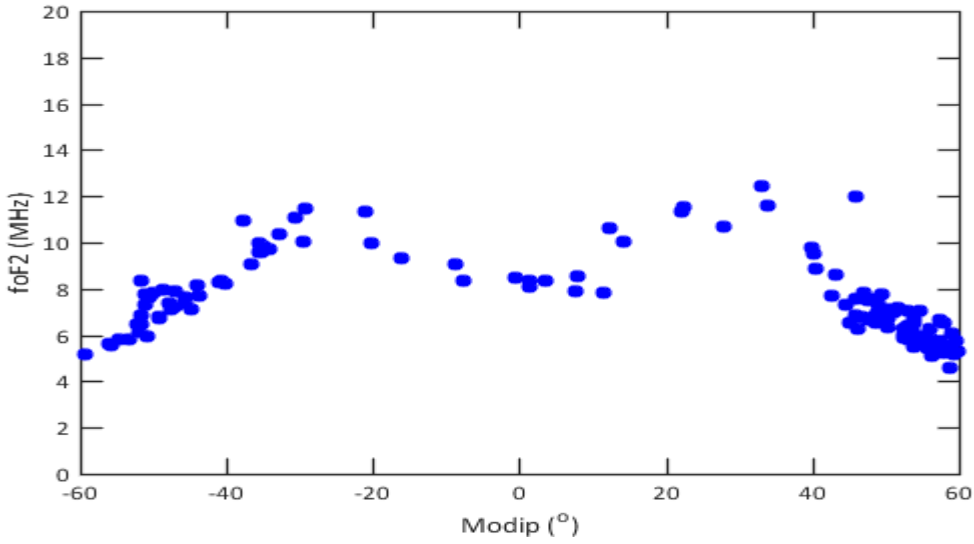

Figure 4.9: Monthly median foF2 observations (blue dots) from all sites at local noon during March, $S_{12} = 10$.

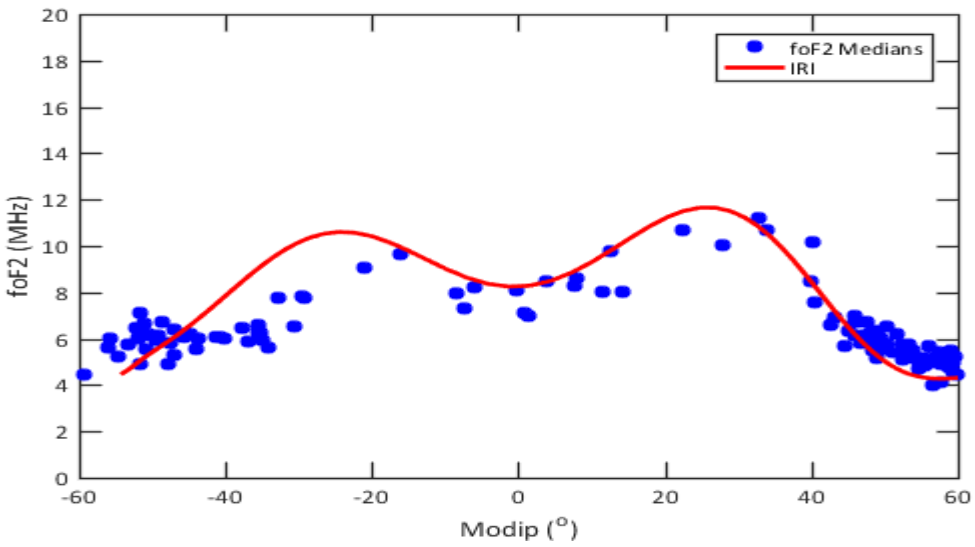

Figure 4.10: IRI's foF2 model (red line) compared to monthly median foF2 data (blue dots) during May at 15 LT for $S_{12} = 10$.

Figures 4.10 to 4.13 display IRI foF2 and the monthly median foF2 data for May at 15 LT as a function of modip at various points in the solar cycle ($S_{12} = 10$, 30, 115, 200). These figures suggest that IRI overestimates foF2 at the southern equatorial anomaly during winter. This does not appear to be as significant in the northern equatorial anomaly. In general it appears that IRI performs better at all latitudes in the Northern Hemisphere than in the Southern Hemisphere. This may be due to more foF2 observations from the Northern Hemisphere contributing to IRI.

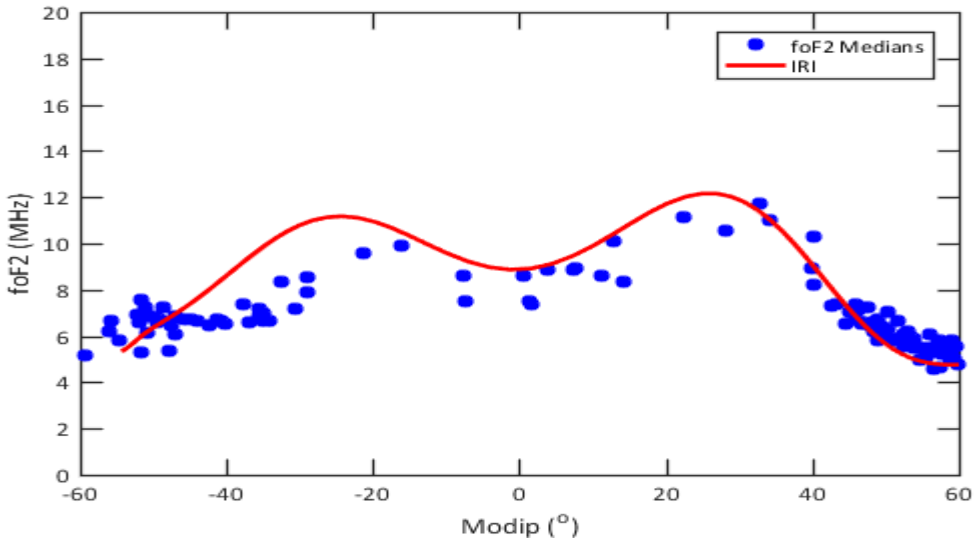

Figure 4.11: IRI's foF2 model (red line) compared to monthly median foF2 data (blue dots) during May at 15 LT for $S_{12} = 30$.

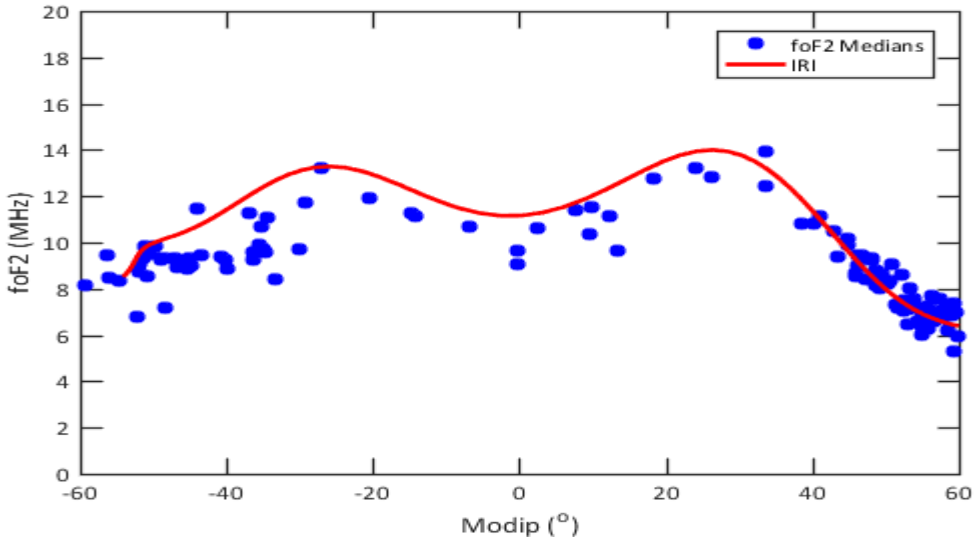

Figure 4.12: IRI's foF2 model (red line) compared to monthly median foF2 data (blue dots) during May at 15 LT for $S_{12} = 115$.

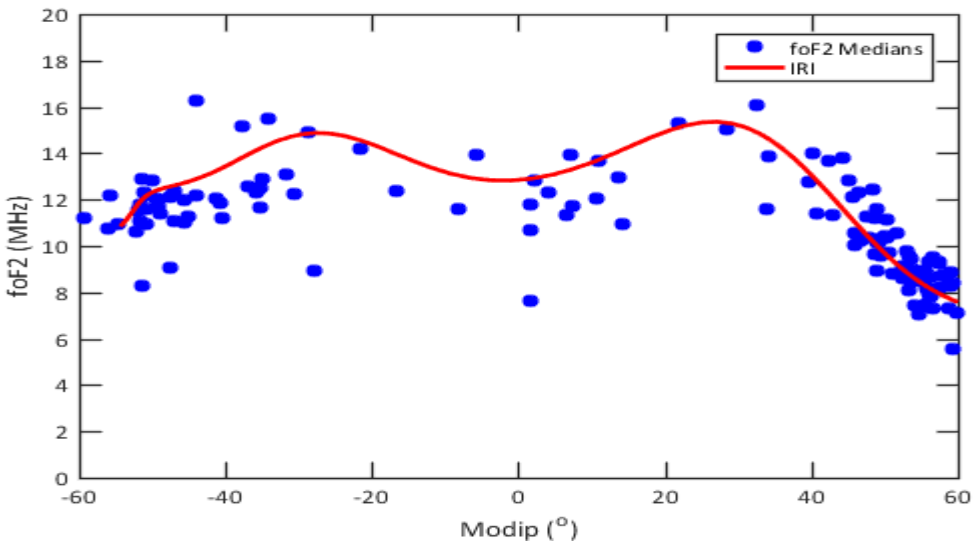

Figure 4.13: IRI's foF2 model (red line) compared to monthly median foF2 data (blue dots) during May at 15 LT for $S_{12} = 200$.

An alternative method of examining the global structure of errors in IRI is to generate plots of the mean foF2 residuals as a function of season and solar cycle (as for Canberra displayed in Figure 4.6) for several sites and order them by their dip angle. This would show the how the behaviour of IRI's foF2 residuals as a function season and solar cycle varies with geomagnetic latitude. Figure 4.14 displays the mean of IRI's residuals for 9 sites at various magnetic latitudes and Figure 4.15 shows the standard deviation of the residuals.

Examination of the mean of the foF2 residuals in Figure 4.14 shows a clear global trend. IRI's foF2 model tends to underestimate at low sunspot numbers and overestimates at high sunspot numbers. It appears IRI's foF2 model does best at the mid latitudes, as shown by Canberra and Juliusruh in the bottom right and top left of the figure respectively. Additionally, IRI does fairly well near the geomagnetic equator, as shown by the middle row of plots, but there is still overestimation at high sunspot numbers, more so than at the mid latitudes, and particularly at Jicamarca. It appears that IRI struggles describing the northern edge of the Northern Hemisphere equatorial anomaly, and the southern edge of the Southern Hemisphere equatorial anomaly (i.e. the outside edges) as seen in Tokyo, Puerto Rico, Port Moresby and Darwin. There also appears to be a seasonal effect in the residuals. Examination of Canberra and Juliusruh shows that IRI overestimates during the Summer, while it underestimates in winter. At Tokyo, Puerto Rico, Port Moresby and Darwin IRI tends to overestimate foF2 during winter and underestimate in summer.

Examination of the standard deviation of the residuals displayed in Figure 4.15 also shows geomagnetic, seasonal and solar cycle influences on the residuals. As

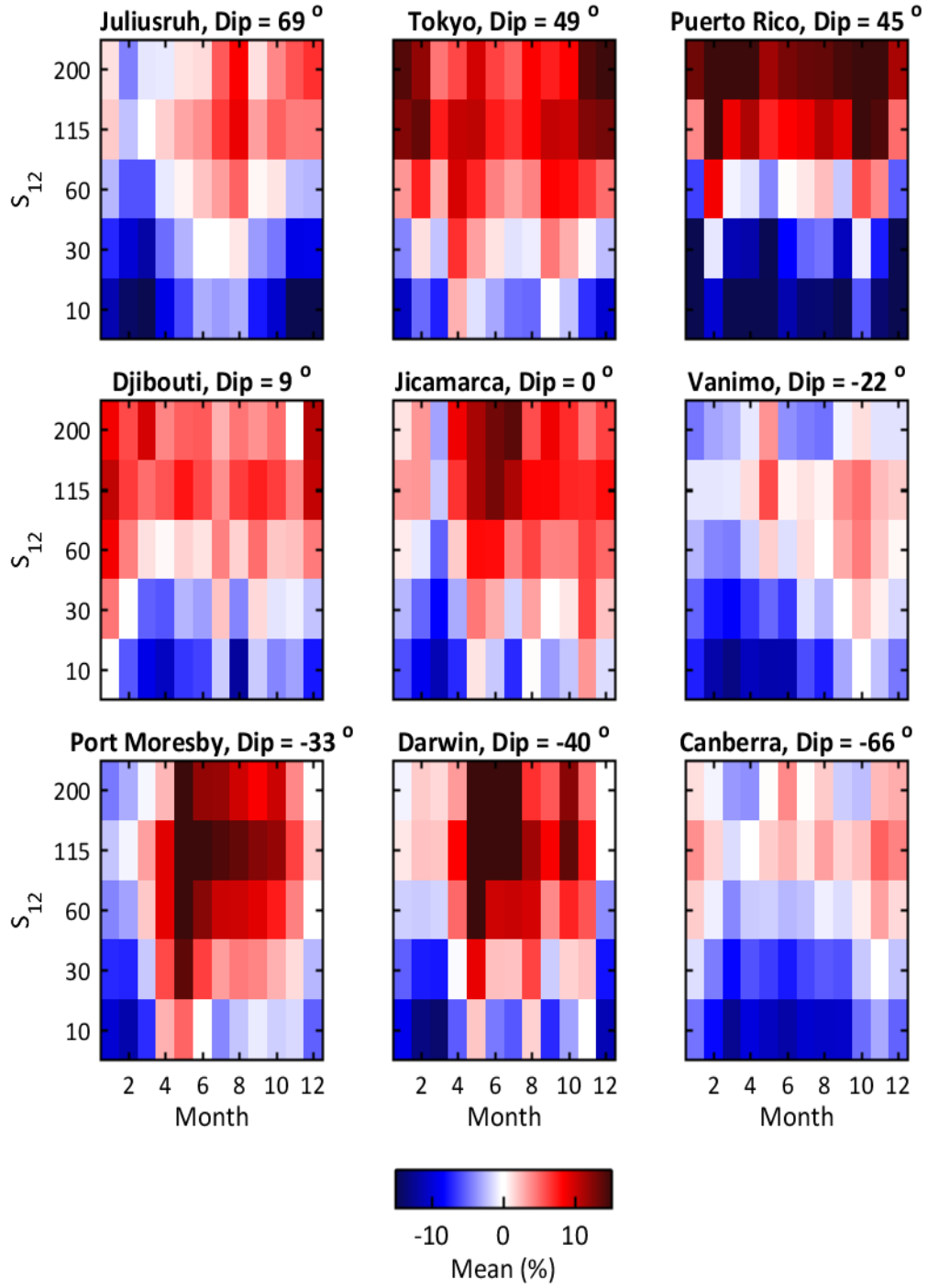

Figure 4.14: The mean foF2 residual between IRI's foF2 model and monthly median observations. From left to right, Top: Juliusruh (Dip = 69 °), Tokyo (Dip = 49 °), Puerto Rico (Dip = 46 °); middle: Djibouti (Dip = 11 °) , Jicamarca (Dip = 1 °), Vanimo (Dip = -22 °); bottom: Port Moresby (Dip = -33 °), Darwin (Dip = -40 °), Canberra (Dip = -66 °) ordered by geomagnetic dip angle.

above, IRI's foF2 model performs best in the mid latitudes at Canberra and Juliusruh. The largest standard deviations occur during winter but they still tend to be lower than at other sites. Tokyo has a similar standard deviation structure to Juliusruh but the standard deviations are larger. The standard deviations tend to be largest at Djibouti, Jicamarca, Vanimo, Port Moresby and Darwin. Overall, it appears that the Southern Hemisphere standard deviations are larger than the Northern Hemisphere.

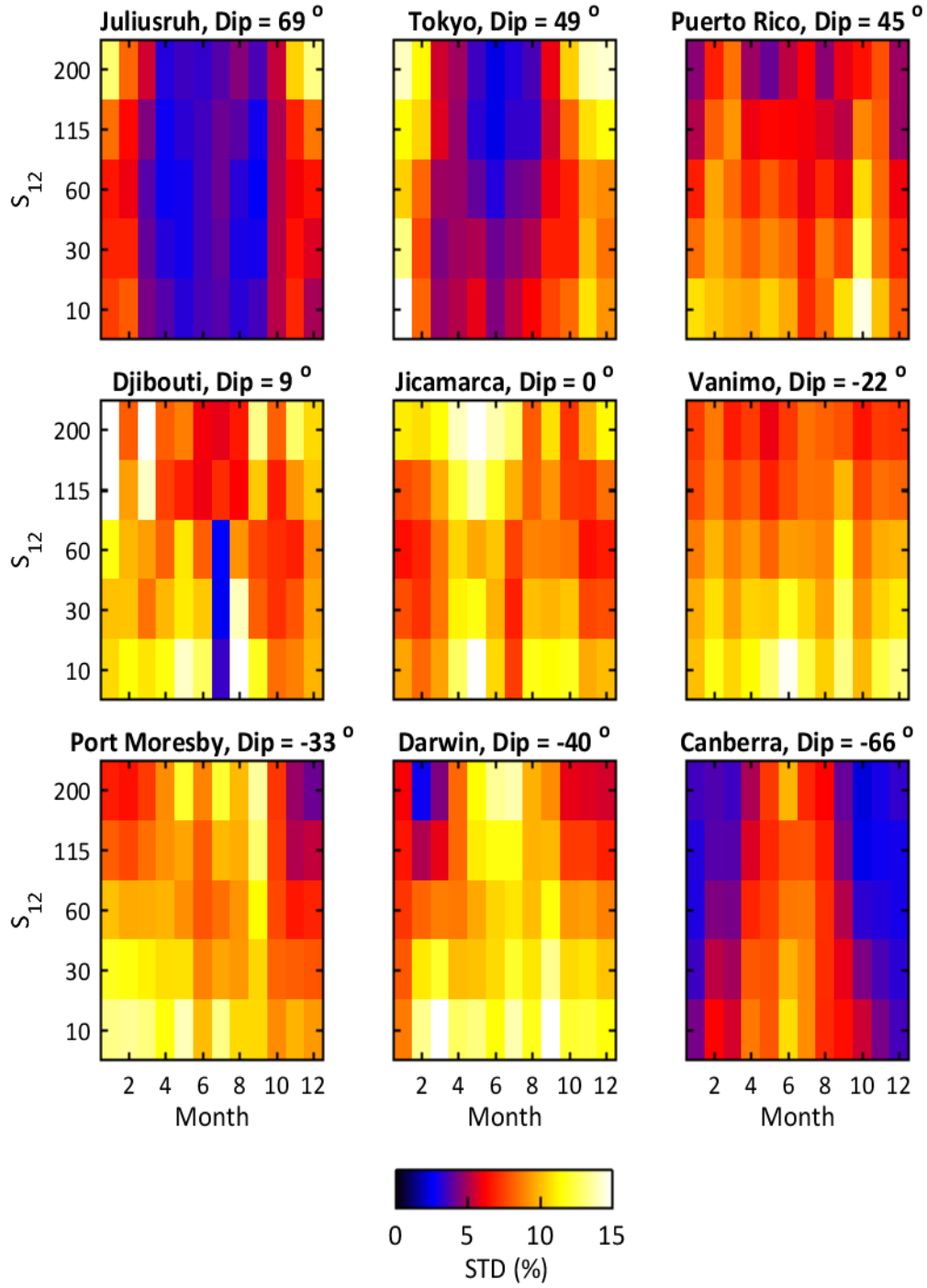

Figure 4.15: The standard deviation of the foF2 residual between IRI's foF2 model and monthly median observations. From left to right, Top: Juliusruh (Dip = 69 °), Tokyo (Dip = 49 °), Puerto Rico (Dip = 46 °); middle: Djibouti (Dip = 11 °) , Jicamarca (Dip = 1 °), Vanimo (Dip = -22 °); bottom: Port Moresby (Dip = -33 °), Darwin (Dip = -40 °), Canberra (Dip = -66 °) ordered by geomagnetic dip angle.

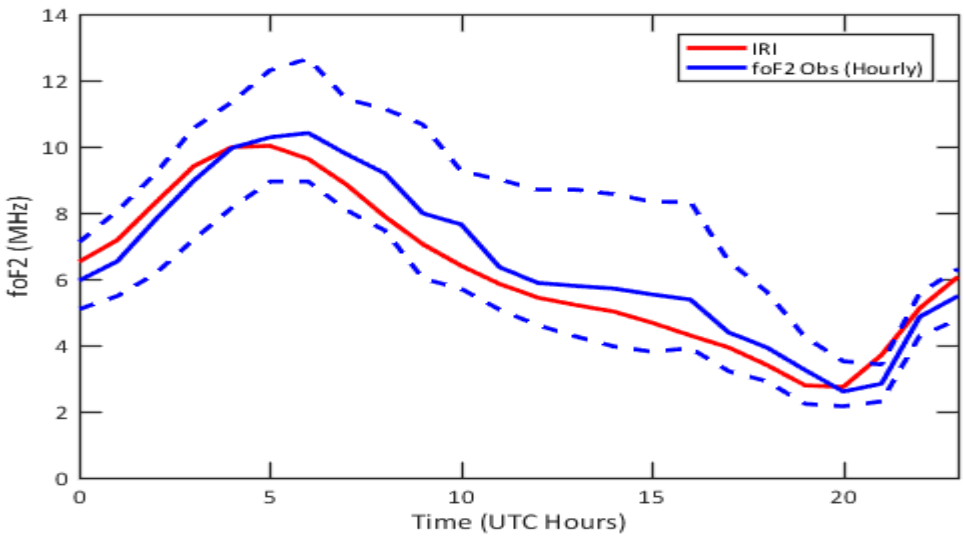

Figure 4.16: A comparison between IRI (red line) and GIRO's monthly median foF2 (blue line) and the tenth and ninetieth percentile (dashed blue lines) at Darwin in February 2017.

4.3 Diurnal Behaviour of IRI

Examination of IRI's foF2 prediction of monthly median hourly GIRO observations for Darwin in February 2017 is displayed in Figure 4.16. The GIRO data was binned hourly to remain consistent with the WDC temporal resolution. Figure 4.16 suggests that IRI leads the observations by ~ 1 hour compared to observations. However, for the rest of this analysis it will be described as a lag. The cross correlation function between the IRI foF2 predictions and the observations was calculated to determine the IRI lag as shown in Figure 4.17. A Gaussian fit was applied to the cross correlation data to obtain sub hourly resolution with a calculated lag at 51 ($\pm$ 4) minutes. To test the Gaussian fit reliability, it was reperformed on GIRO Darwin data which had been binned at 15 minute resolution (displayed in Figure 4.18), this produced a lag of 48 ($\pm$ 0.3) minutes. These two lags agree to $\sim$ 3 minutes and the error bars overlap, which validates the lag analysis of the lower temporal resolution data.

Examination of Figures 4.19, 4.20 and 4.21, for Canberra, Darwin and Vanimo respectively at different months and years also shows a temporal offset of the IRI predictions of foF2.

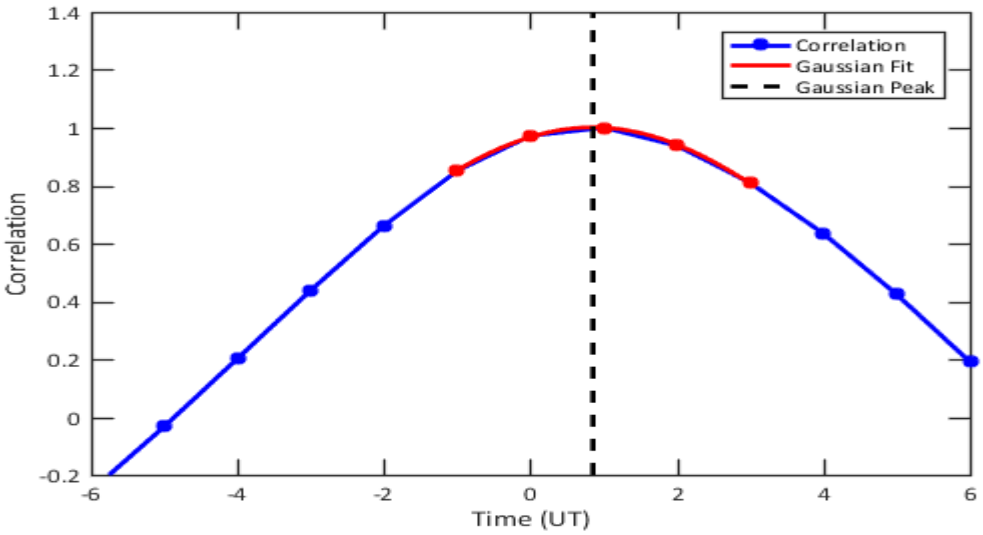

Figure 4.17: The cross correlation between IRI and Darwin monthly median foF2 observation in February 2017.

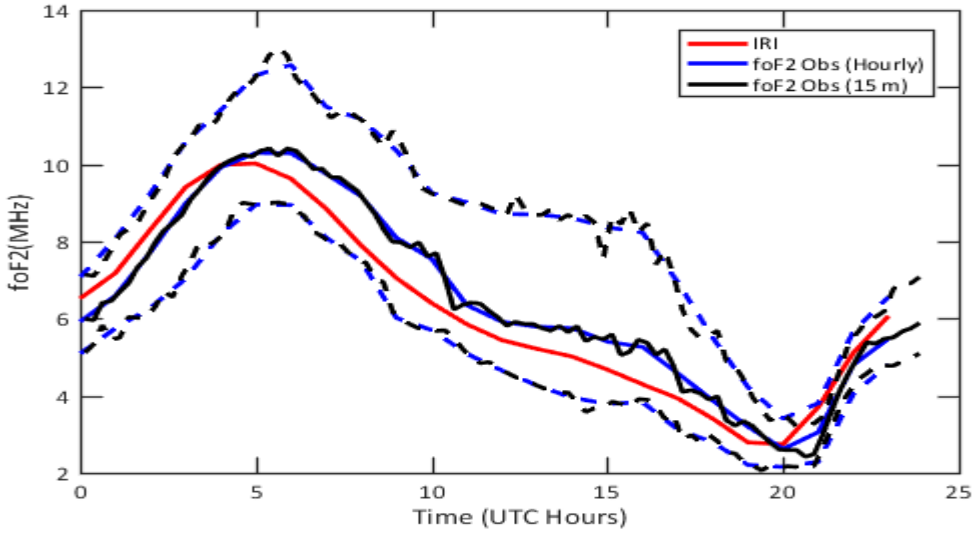

Figure 4.18: A comparison between IRI's monthly median foF2 prediction (red line) and monthly median foF2 observation and the tenth and ninetieth percentile (dashed lines) at hourly (blue) and 15 minute (black) resolution at Darwin in February 2017.

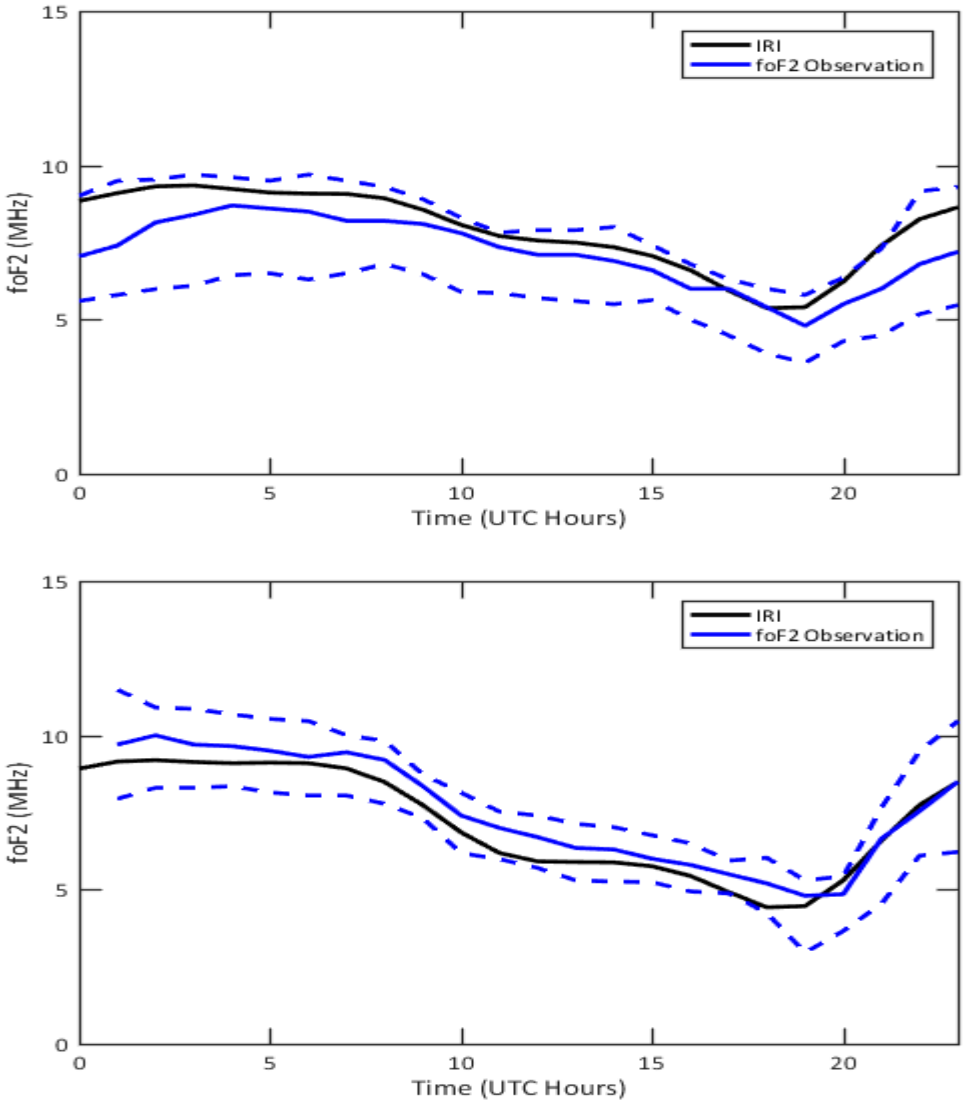

Figure 4.19: A comparison between IRI (black line) and monthly median foF2 (blue line) and the tenth and ninetieth percentile (dashed blue lines) at Canberra in February 1990 (top) and March 1999 (bottom).

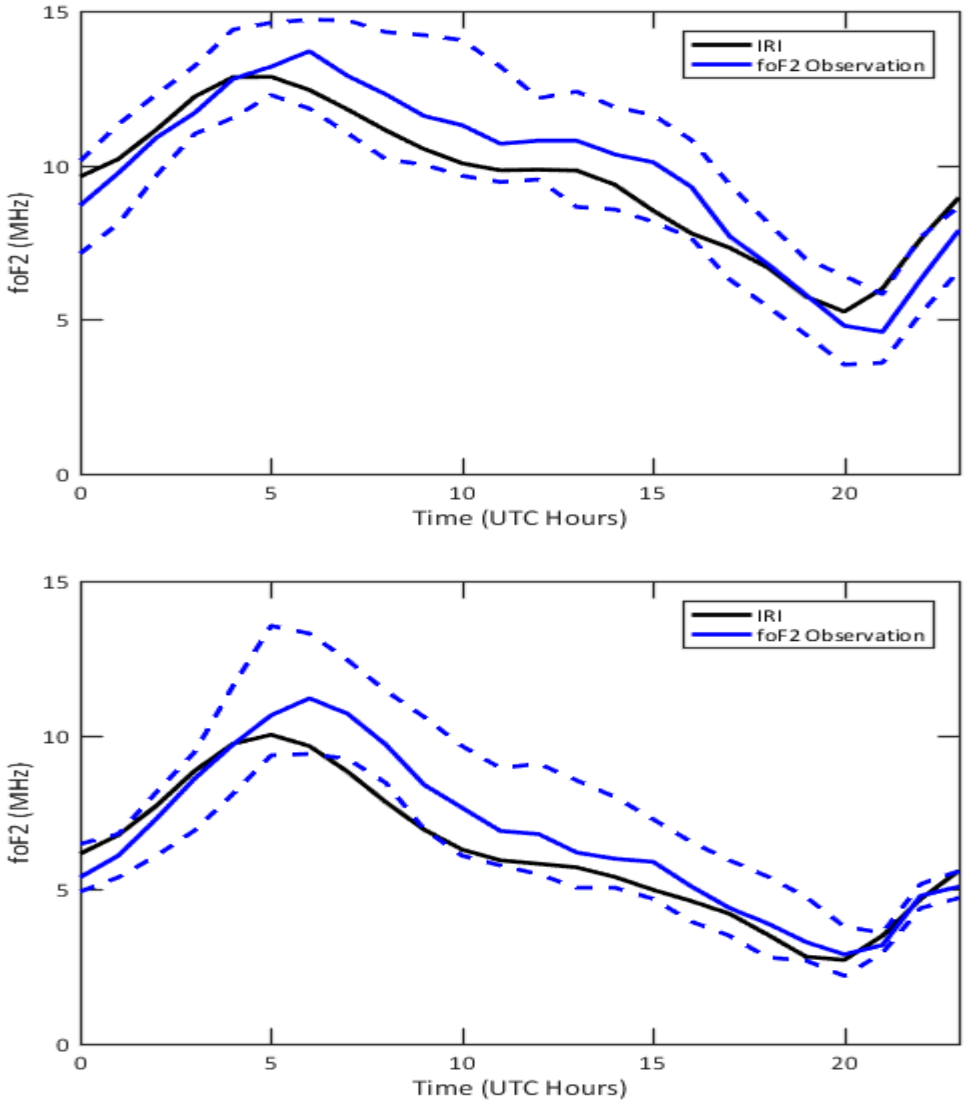

Figure 4.20: A comparison between IRI (black line) and monthly median foF2 (blue line) and the tenth and ninetieth percentile (dashed blue lines) at Darwin in February 1999 (top) and January 2010 (bottom).

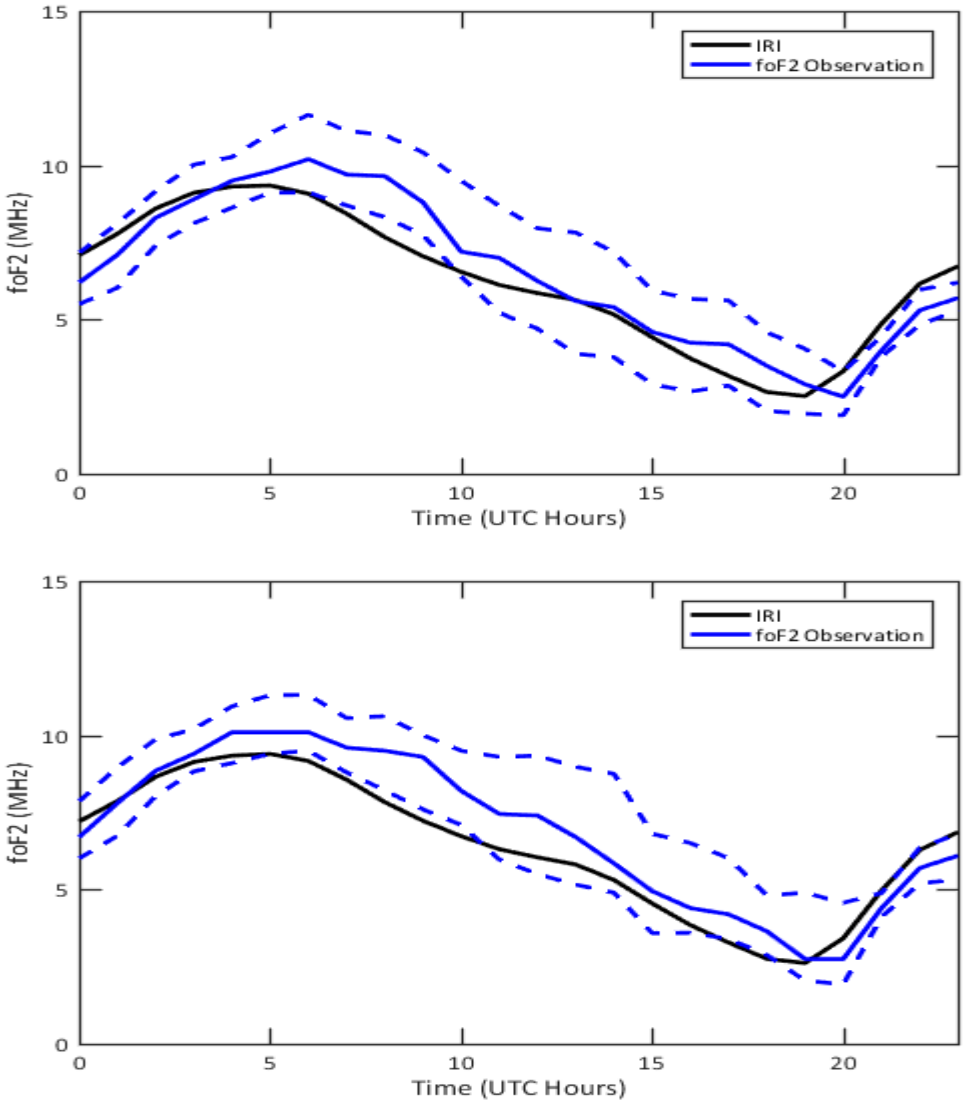

Figure 4.21: A comparison of foF2 between IRI (black line) and monthly median (blue line) and the tenth and ninetieth percentile (dashed blue lines) estimates at Vanimo in February 1996 (top) and February 2007 (bottom).

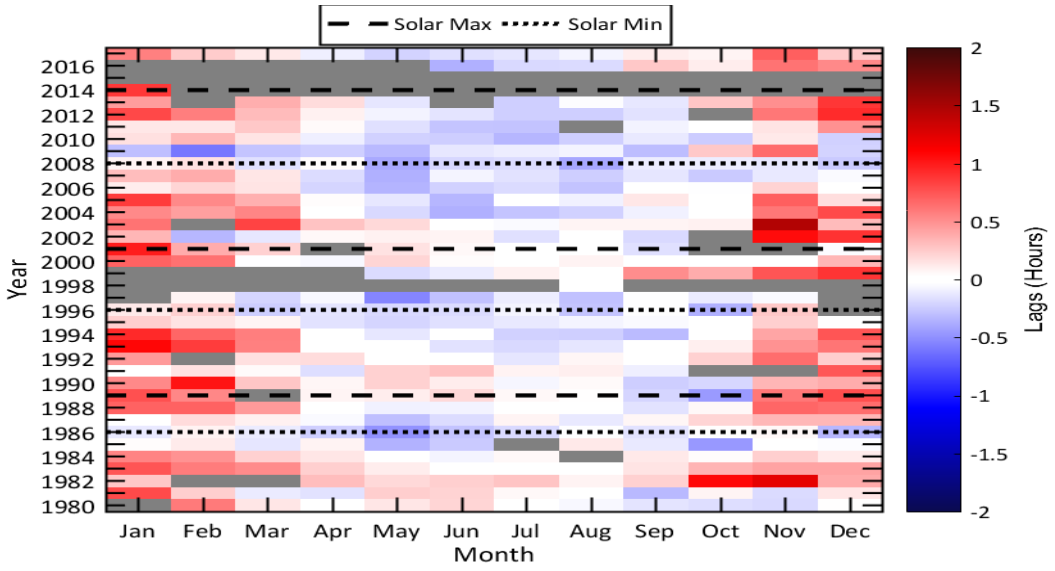

Figure 4.22: Estimates of the lag in IRI2016 foF2 obtained using measurements from Canberra as a function of season and year. It also indicates solar maximum years (long dash line) and solar minimum years (short dash line).

The IRI cross correlation lag was calculated at Canberra for each year and month of available foF2 observations and are displayed in Figure 4.22. These lag data were binned according to S_{12} and displayed in Figure 4.23. Clearly, IRI tends to have a lag of ~ 0.5 hours in summer becoming larger as S_{12} increases. During winter, the lag is negative becoming worse as S_{12} decreases.

The lag analysis was repeated at Darwin, and is displayed in Figure 4.24. Overall the lag tends to be larger at Darwin than Canberra. In addition, the seasonal and solar cycle dependence is quite different. For instance, during the summer months, the lag increases with decreasing S_{12}, the opposite is noted at Canberra.

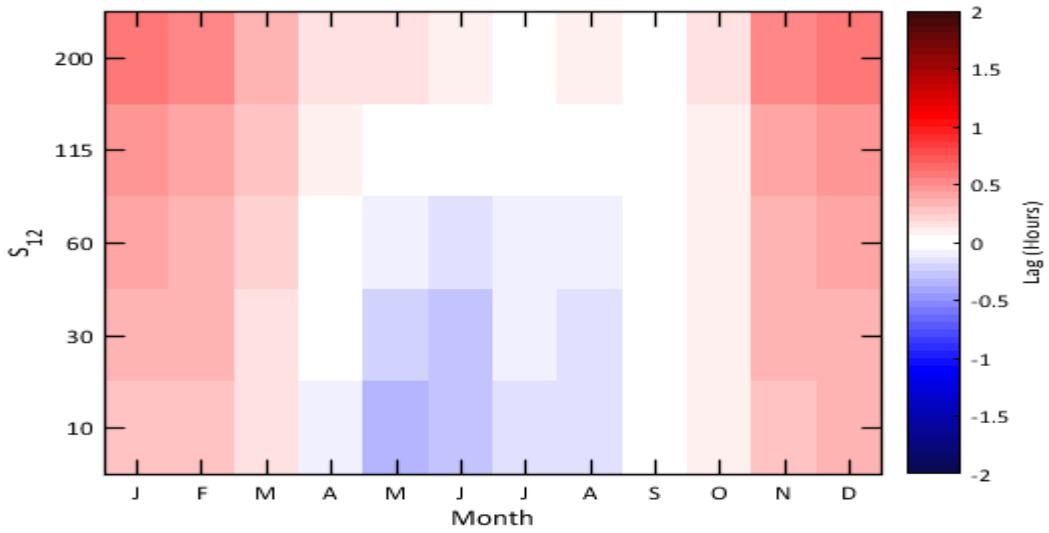

Figure 4.23: Estimates of the lag in IRI2016 foF2 at Canberra as a function of season and sunspot number, S_{12}.

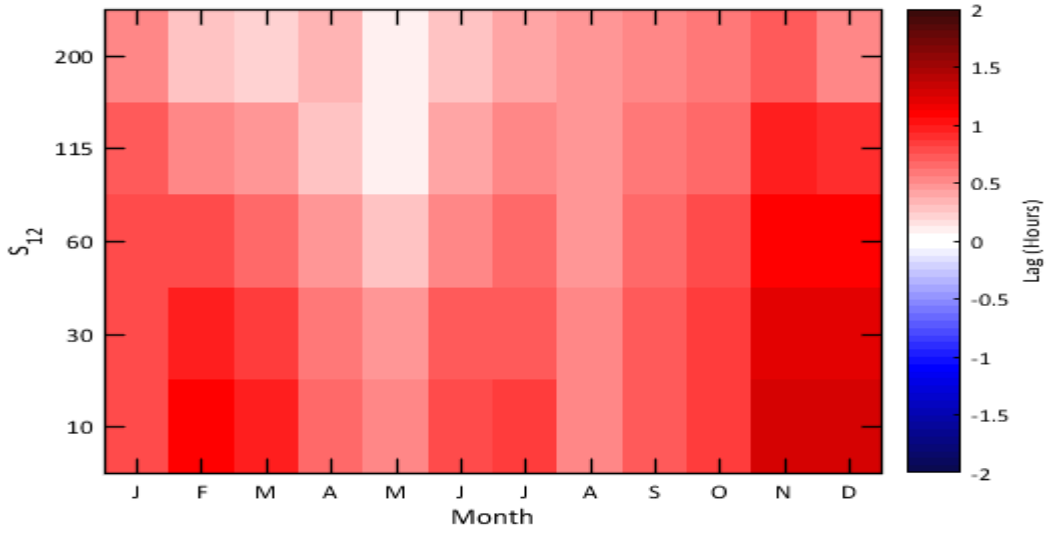

Figure 4.24: Estimates of the lag in IRI2016 foF2 at Darwin as a function of season and sunspot number, S_{12}.

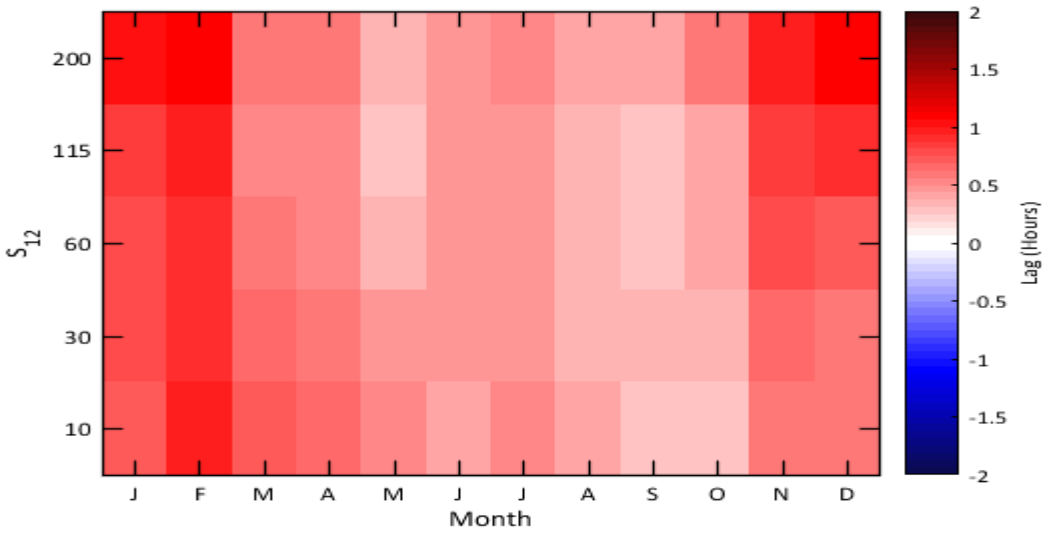

Figure 4.25: Estimates of the lag in IRI2016 foF2 at Vanimo as a function of season and sunspot number, S_{12}.

The cross correlation lag results at Vanimo are displayed in Figure 4.25. Again the dependence of the lag on season and S_{12} is different to Darwin. At Vanimo the lag is relatively constant with S_{12}, but there is a strong seasonal dependence with larger lags during summer. Canberra, Darwin and Vanimo share some similarities as summer tends to have a larger lag than winter and the lags are typically between 0.5 hours to 1 hour.

The lag results from Darwin and Vanimo shows interesting results. The shoulder (outer edge) of the southern equatorial anomaly appears to have greater variation at solar minimum than at solar maximum, but the peak (Vanimo) is more evenly dynamic across the solar cycle but perhaps a little more dynamic at solar maximum.

Finally, an example of the foF2 lag analysis at a Northern Hemisphere site, Juliusruh, is displayed in Figure 4.26. The lag is significantly less than at Canberra a comparable Southern hemisphere mid latitude site. IRI exhibits only a small lag at Juliusruh being slightly positive in winter and a much smaller negative lag in summer.

To determine the global structure of the foF2 lags, a few selected sites were sorted by dip as illustrated in Figure 4.27. There are positive lags in the low latitudes as shown by the Darwin, Port Moresby, Vanimo, Jicamarca plots. In all these plots, the lag appears to be larger in October to February and lower in March to August. Examination of the mid latitude sites, Juliusruh, Tokyo and Canberra, indicates that there tends to be slight negative lags in winter, much more so than at the

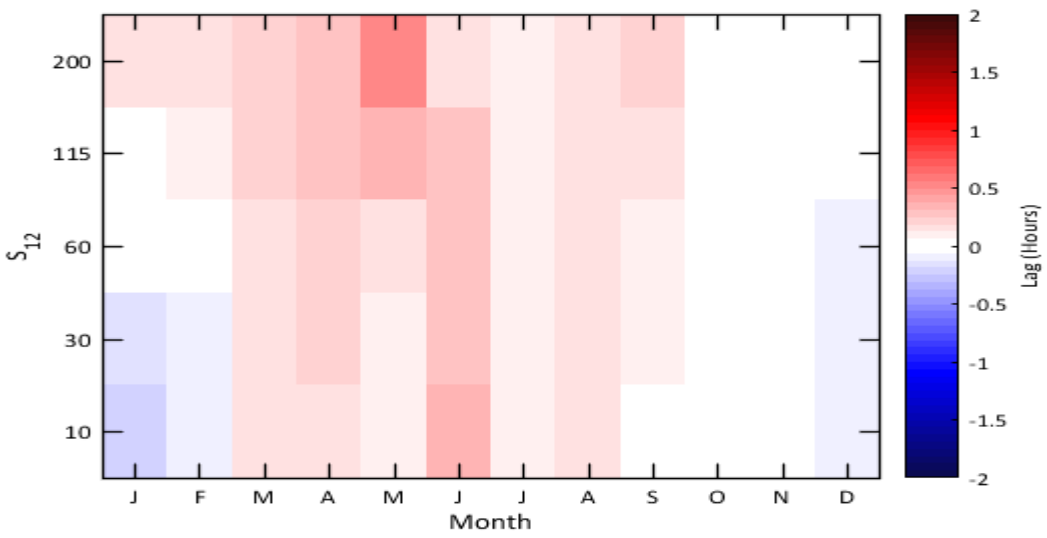

Figure 4.26: Estimates of the lag in IRI2016 foF2 obtained using measurements from Juliusruh as a function of sunspot number, S_{12}.

other sites. At the geomagnetic equator the lags are small and are similar to those estimated at mid latitude sites Juliusruh and Canberra.

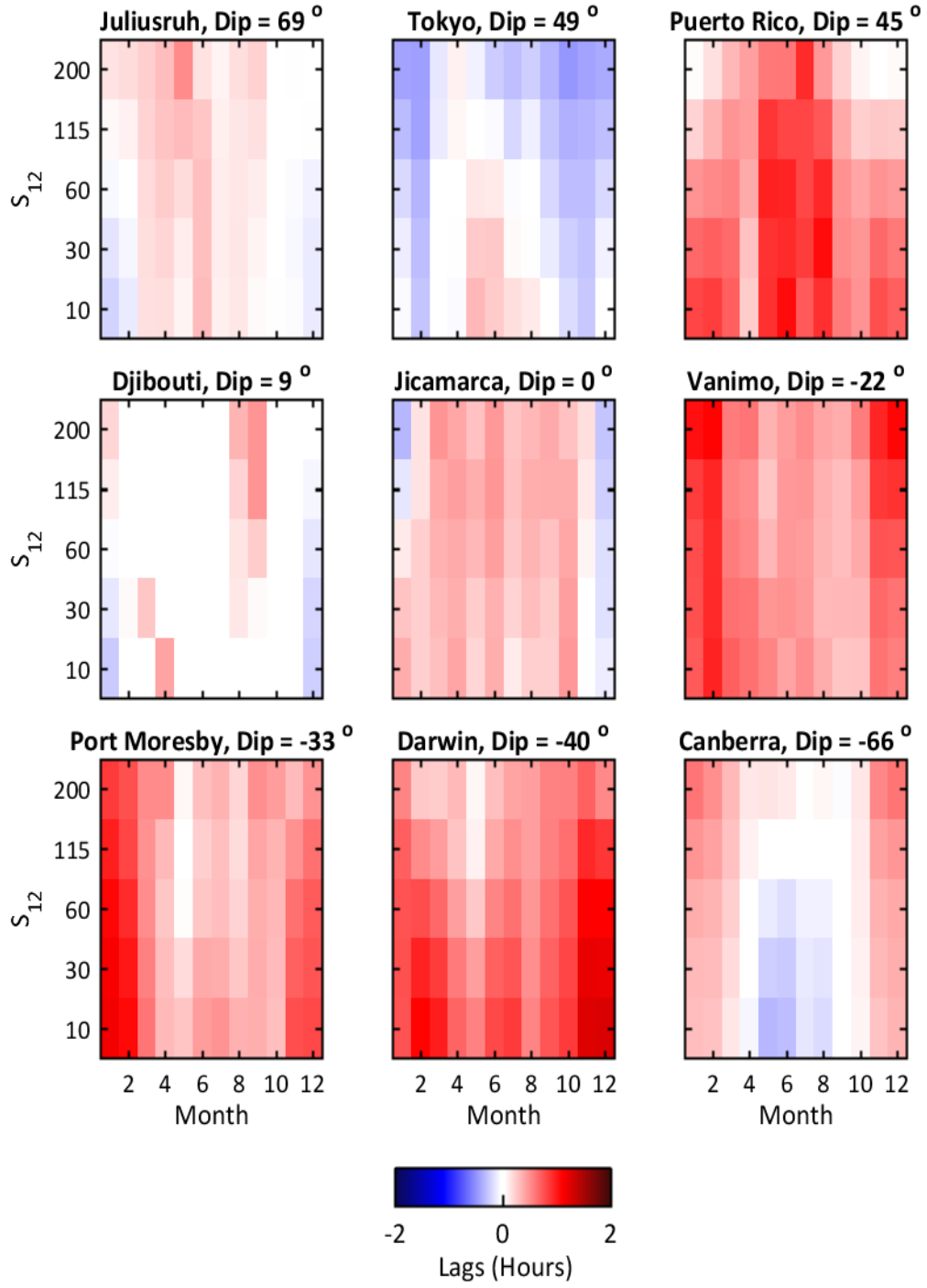

Figure 4.27: The cross correlation lags in foF2 between IRI's foF2 model and monthly median observations. From left to right, Top: Juliusruh (Dip = 69 °), Tokyo (Dip = 49 °), Puerto Rico (Dip = 46 °); middle: Djibouti (Dip = 11 °), Jicamarca (Dip = 1 °), Vanimo (Dip = -22 °); bottom: Port Moresby (Dip = -33 °), Darwin (Dip = -40 °), Canberra (Dip = -66 °) ordered by geomagnetic dip angle.

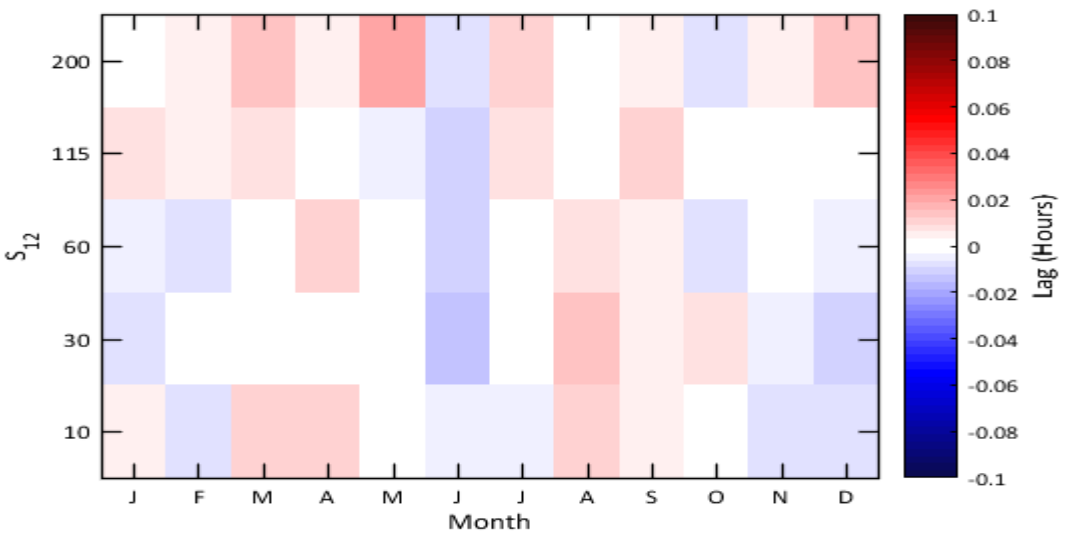

Figure 4.28: Estimates of the lag in foF2 in Lag Corrected IRI2016 at Vanimo. Note the change in scale compared to the previous lag plots.

4.4 Lag Corrected IRI (LCIRI)

Given the observed lag in IRI's foF2 model it is possible to use the lag estimate to produce a "lag corrected" IRI (LCIRI) which may improve IRI's foF2 prediction. LCIRI was produced by adding a temporal offset based off the cross correlation results for that particular month, year and site. The cross correlation lags were recalculated for LCIRI as illustrated in Figure 4.28 to show the temporal correction worked as expected. Note the change of scale Figure 4.28 compared to the other cross correlation lag figures. The effects of the lag correction on the standard deviation of IRI's foF2 residual is illustrated in Figures 4.29, 4.30 and 4.31 for Canberra, Darwin and Vanimo respectively. There is a clear improvement at Darwin and Vanimo from this lag correction with a significant reduction in the standard deviation. At Canberra there was only a slight improvement which may be due to the magnitude of the IRI lag being less than at Darwin and Vanimo.

The mean and standard deviations of the foF2 residuals at the 9 sites were calculated to determine the improvement from application of the correction. Lag corrected IRI foF2 residual means and standard deviations are shown in Figures 4.32 and 4.33. For comparison, the corresponding uncorrected IRI mean and standard deviation plots are in Figures 4.14 and 4.15. Overall, the lag correction provides little change to the mean residual. However, there are significant reductions to the standard deviation at some sites. While Juliusruh, Tokyo, Canberra and Djibouti do not display much improvement, Puerto Rico, Jicamarca, Vanimo, Port Moresby and Darwin benefit immensely. With the lag correction, these sites perform as well as Juliusruh and Canberra. After the lag correction process there is still a residual

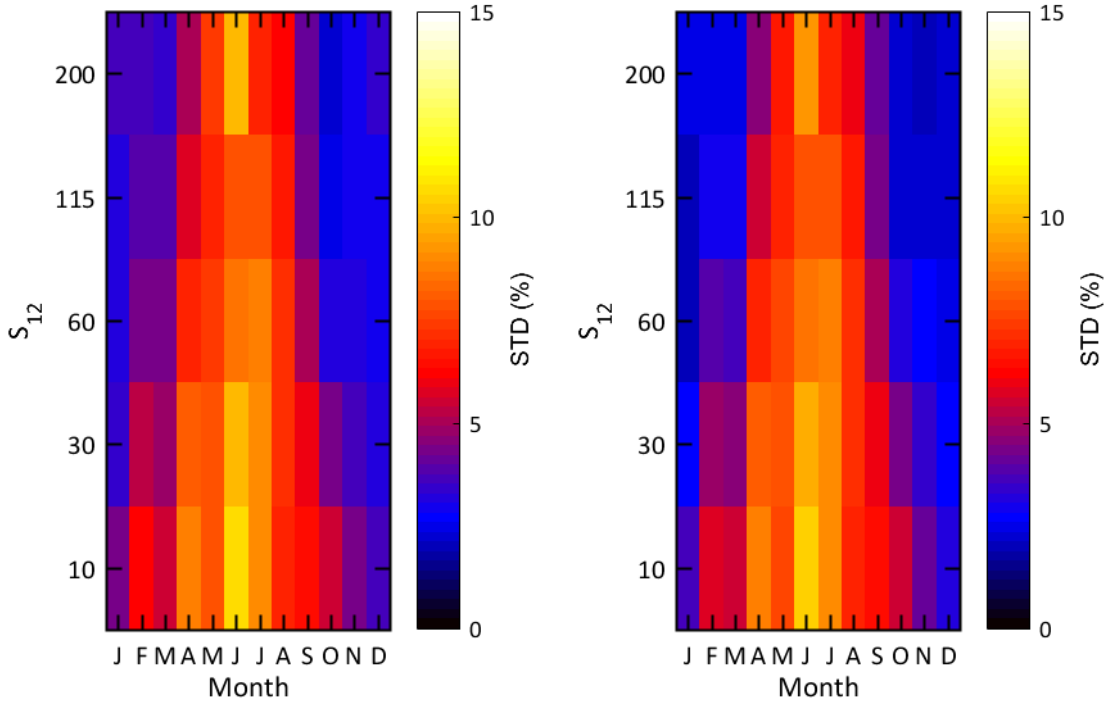

Figure 4.29: A comparison of IRI (left) and lag corrected IRI (right) foF2 residual standard deviations from Canberra.

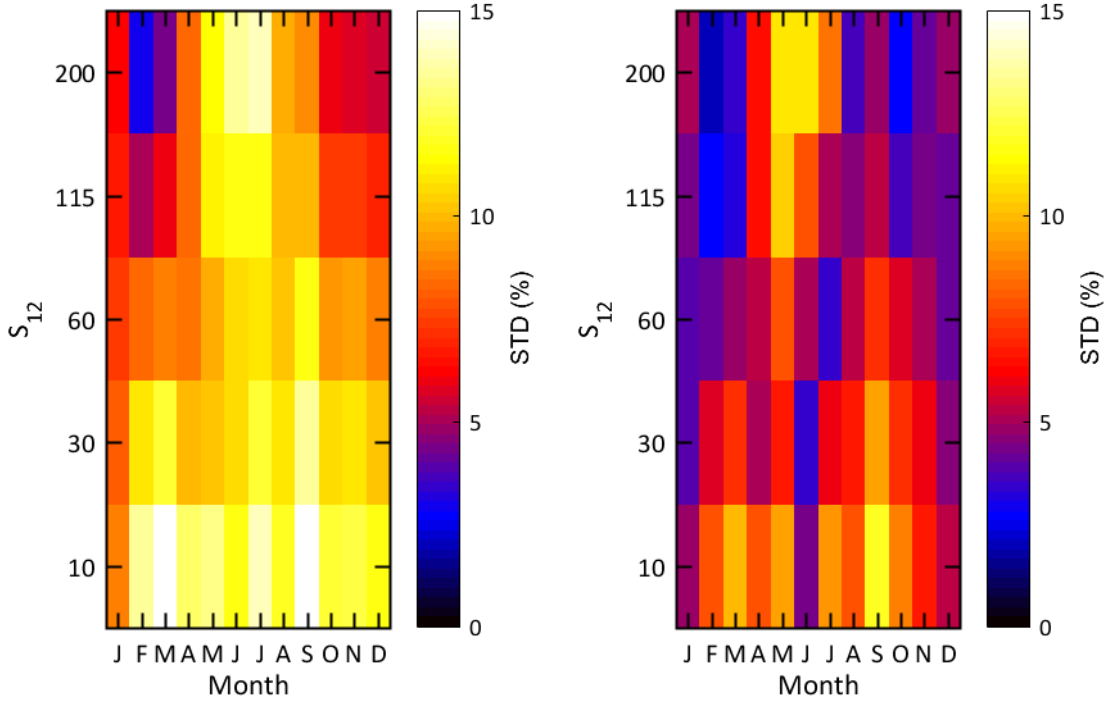

Figure 4.30: A comparison of IRI (left) and lag corrected IRI (right) foF2 residual standard deviations from Darwin.

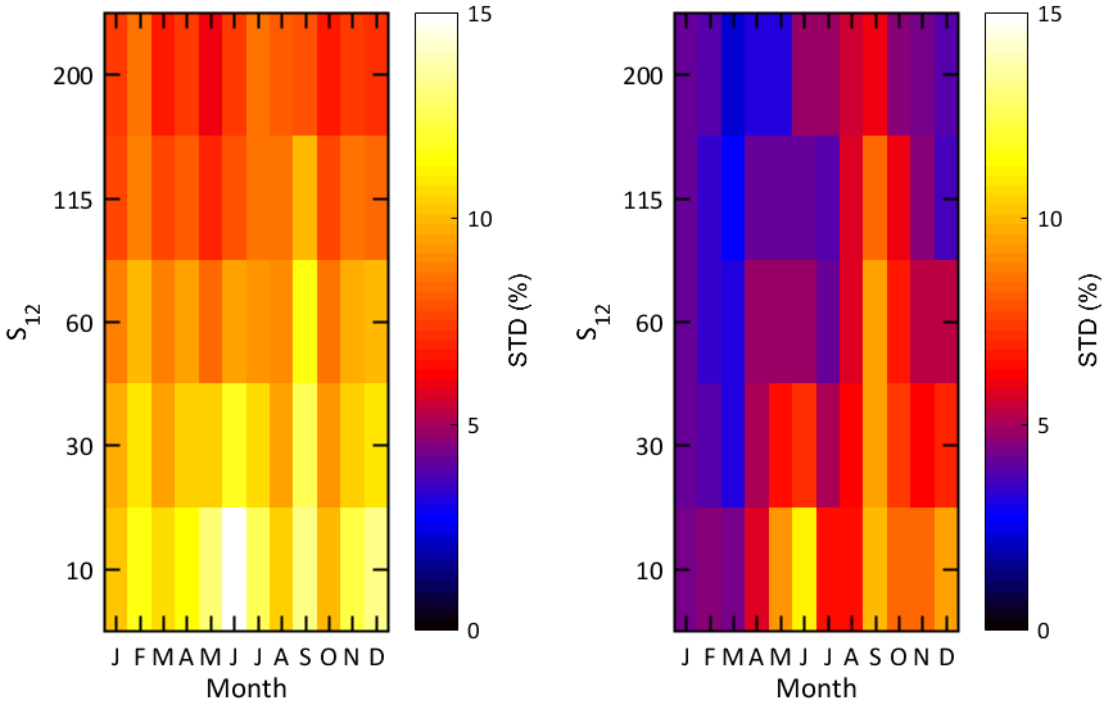

Figure 4.31: A comparison of IRI (left) and lag corrected IRI (right) foF2 residual standard deviations from Vanimo.

present in IRI's foF2 model which is worse during winter solar minimum, suggesting other sources of error.

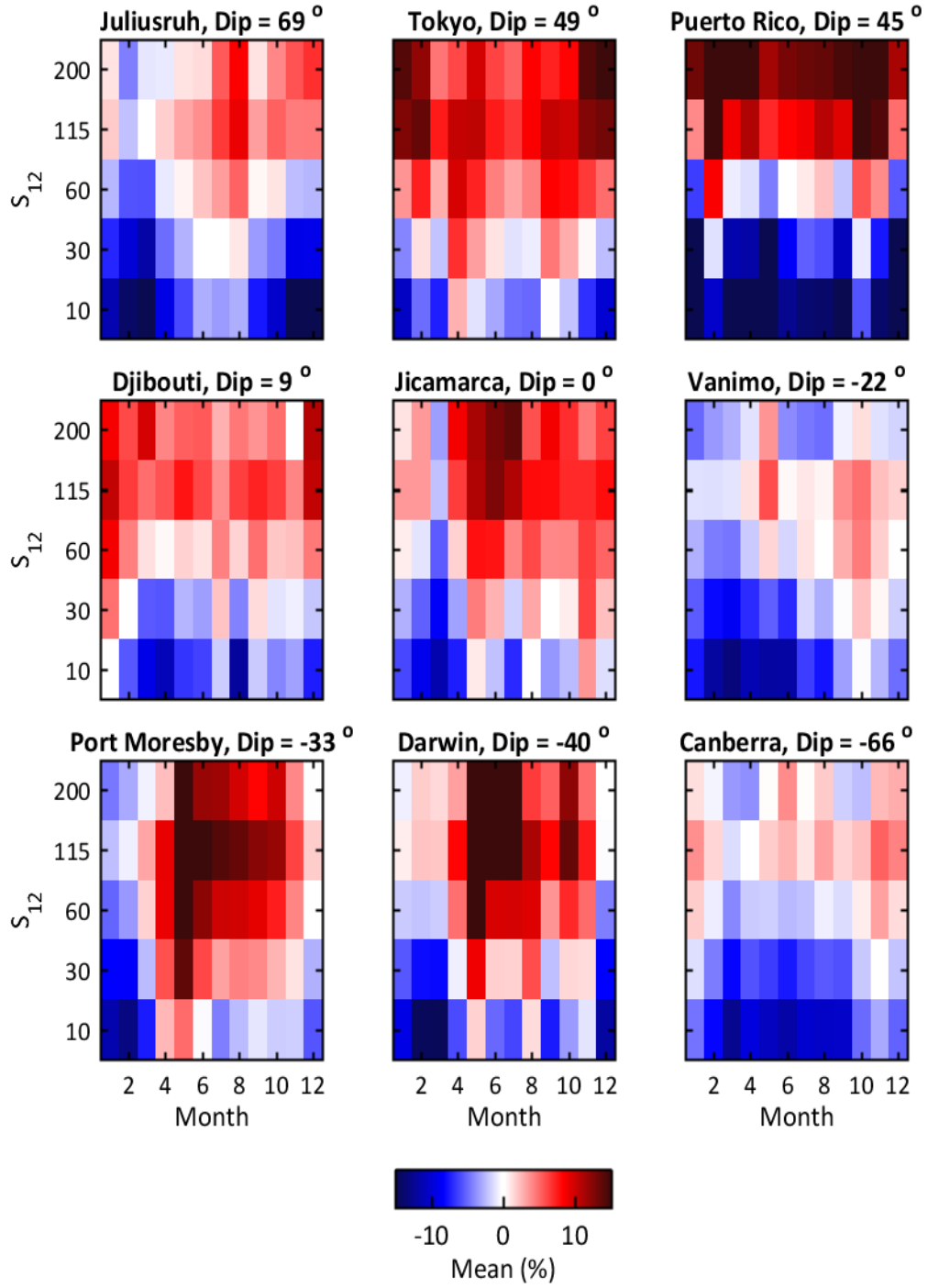

Figure 4.32: The mean foF2 residual between LCIRI's foF2 model and monthly median observations. From left to right, Top: Juliusruh (Dip = 69 °), Tokyo (Dip = 49 °), Puerto Rico (Dip = 46 °); middle: Djibouti (Dip = 11 °) , Jicamarca (Dip = 1 °), Vanimo (Dip = -22 °); bottom: Port Moresby (Dip = -33 °), Darwin (Dip = -40 °), Canberra (Dip = -66 °) ordered by geomagnetic dip angle.

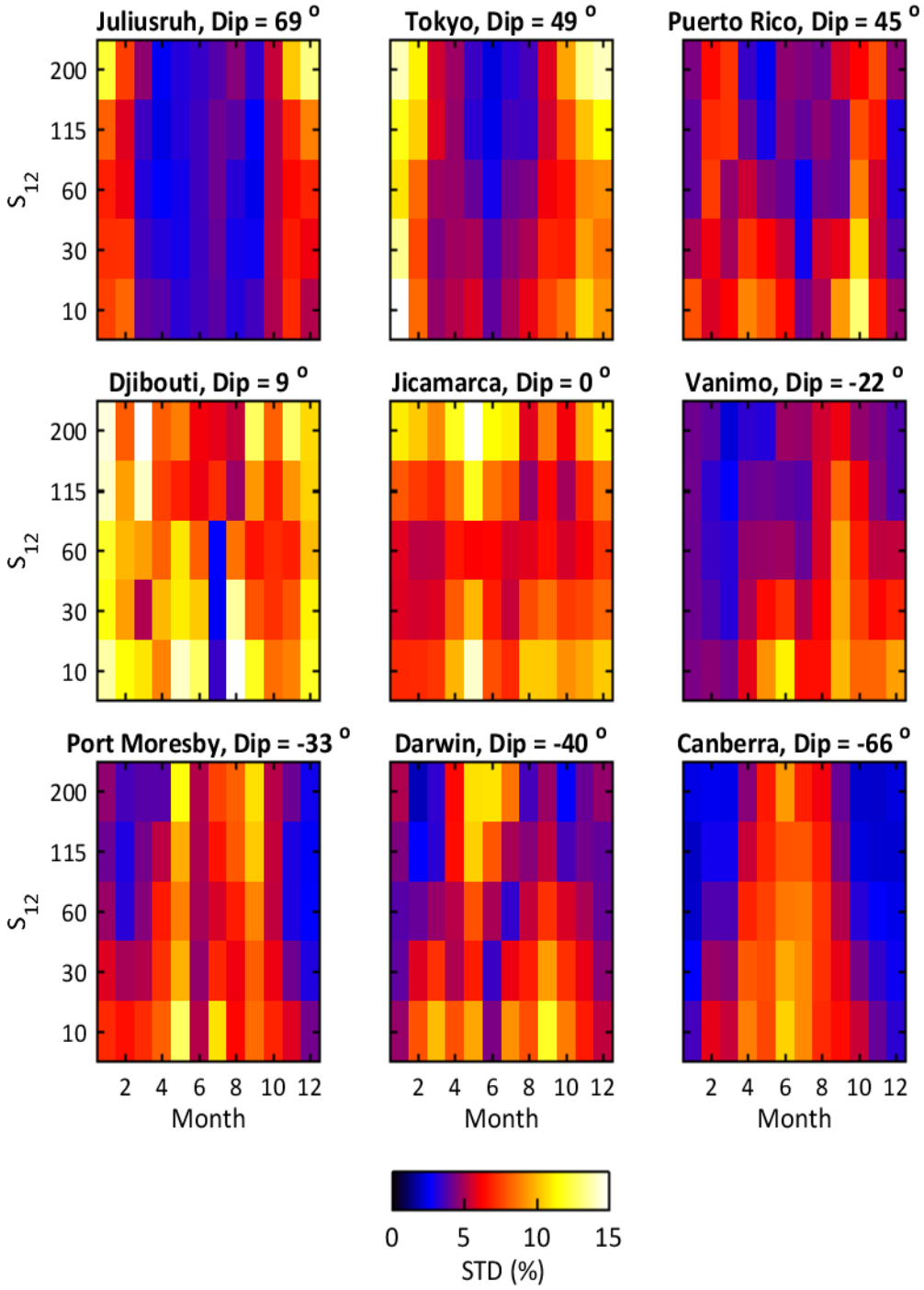

Figure 4.33: The standard deviation of the foF2 residual between LCIRI's foF2 model and monthly median observations. From left to right, Top: Juliusruh (Dip = 69 °), Tokyo (Dip = 49 °), Puerto Rico (Dip = 46 °); middle: Djibouti (Dip = 11 °), Jicamarca (Dip = 1 °), Vanimo (Dip = -22 °); bottom: Port Moresby (Dip = -33 °), Darwin (Dip = -40 °), Canberra (Dip = -66 °) ordered by geomagnetic dip angle.

Site	Dip (°)	RMS Mean (%)	Lag Corrected RMS Mean (%)	RMS STD (%)	Lag Corrected RMS STD(%)
Canberra	-66	5	5	6	5.8
Darwin	-40	8.9	8.9	10.3	6.4
Port Moresby	-33	9.2	9.3	9.7	6.5
Vanimo	-22	4.9	5.1	9.7	5.8
Jicamarca	1	6.4	6.4	10.2	8.5
Djibouti	11	5.6	5.6	10.6	10.5
Puerto Rico	46	12.3	12.4	8	5.9
Tokyo	49	7.7	7.7	7.8	7.7
Juliusruh	69	6.5	6.5	5.6	5.5

Table 4.2: Comparison of IRI and lag corrected IRI foF2 residual mean and standard deviations.

The RMS mean and standard deviation for IRI and Lag Corrected IRI for the nine selected sites are shown in Table 4.2. This table shows that Lag Corrected IRI has a reduced RMS standard deviation at all of the selected sites with the greatest improvement at Darwin, where the RMS standard deviation is reduced from 10.3 % to 6.4 %. The RMS mean does not appear to be affected by the lag correction, as suggested by comparing the mean foF2 residuals in Figures 4.14 and 4.32.

The analysis so far has focused on only 9 sites. Figures 4.34 to 4.36 display histograms of the RMS mean and standard deviation foF2 residuals for LCIRI and IRI utilizing all available sites grouped into different geomagnetic regions. These regions are the Northern Hemisphere mid latitudes (74 contributing sites), low latitudes (50 contributing sites) and Southern Hemisphere mid latitudes (36 contributing sites). Figure 4.34 shows that the mean appears to be slightly worse after the lag correction, but the standard deviation had some improvements as LCIRI lower bins were more populated. At low latitudes while the lag correction provided little improvement to the mean residual there was a large improvement to the standard deviation. At the Southern Hemisphere mid latitudes the lag correction does not provide much improvement to the RMS mean but with some improvement to the RMS standard deviation. Overall, it appears the lag correction can provide an improvement to the precision of IRI's foF2 model in the low latitudes but not at the mid latitudes.

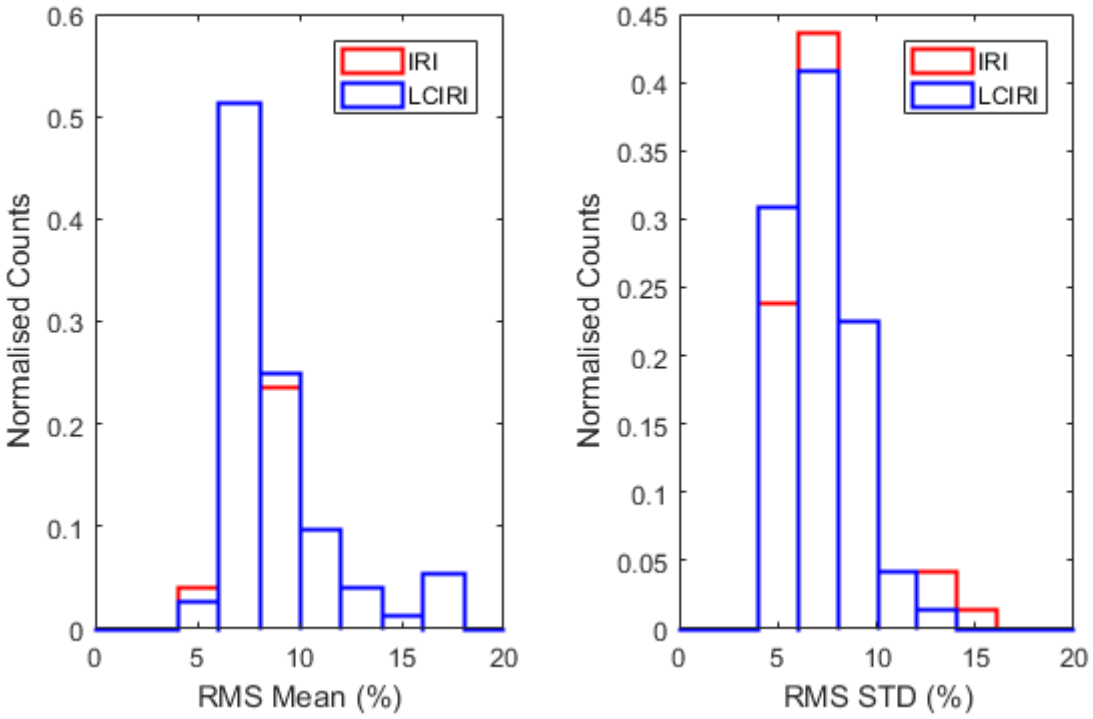

Figure 4.34: RMS of the mean (left) and standard deviation (right) of IRI's (red) and LCIRI's (blue) foF2 residuals calculated for the northern mid latitudes (45 ° Dip to 70 ° Dip) using 74 available sites.

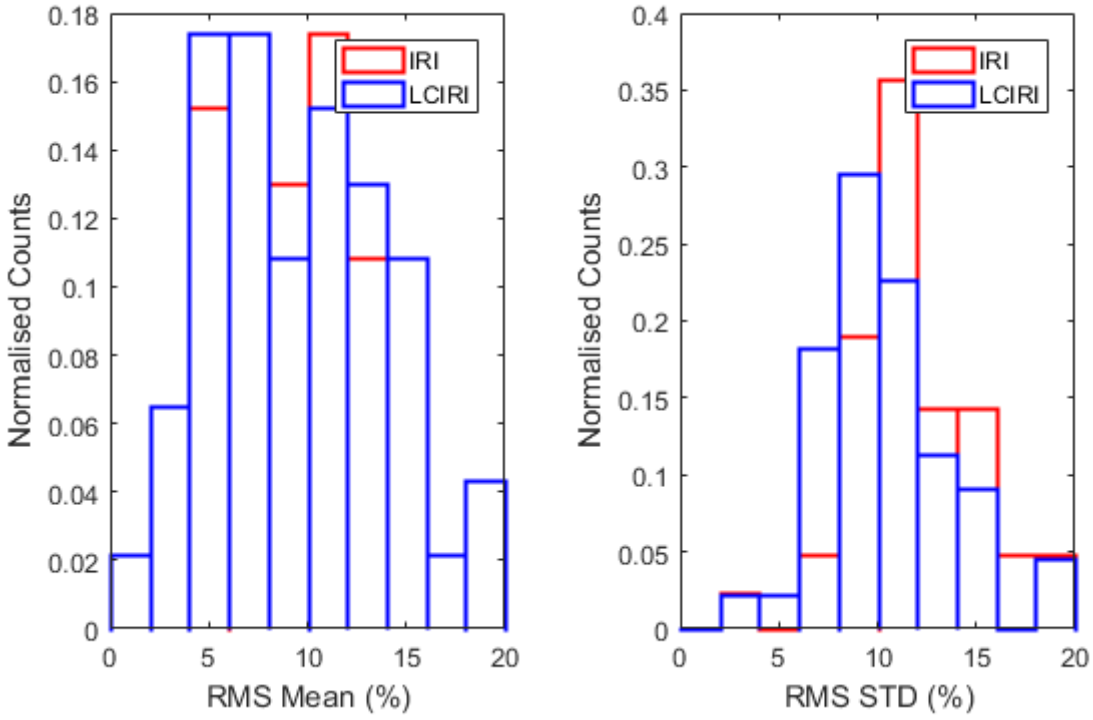

Figure 4.35: RMS of the mean (left) and standard deviation (right) of IRI's (red) and LCIRI's (blue) foF2 residuals calculated for the low latitudes (-45 ° Dip to 45 ° Dip) using 50 available sites.

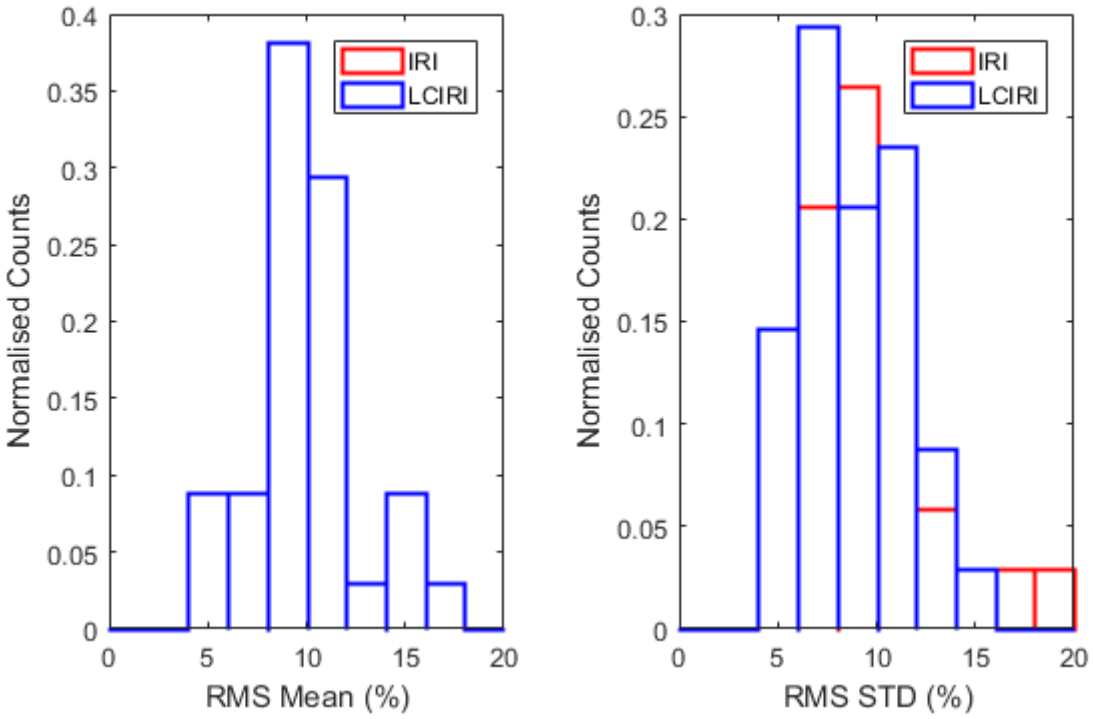

Figure 4.36: RMS of the mean (left) and standard deviation (right) of IRI's (red) and LCIRI's (blue) foF2 residuals calculated for the southern mid latitudes (-45 ° Dip to -70 ° Dip) using 36 available sites.

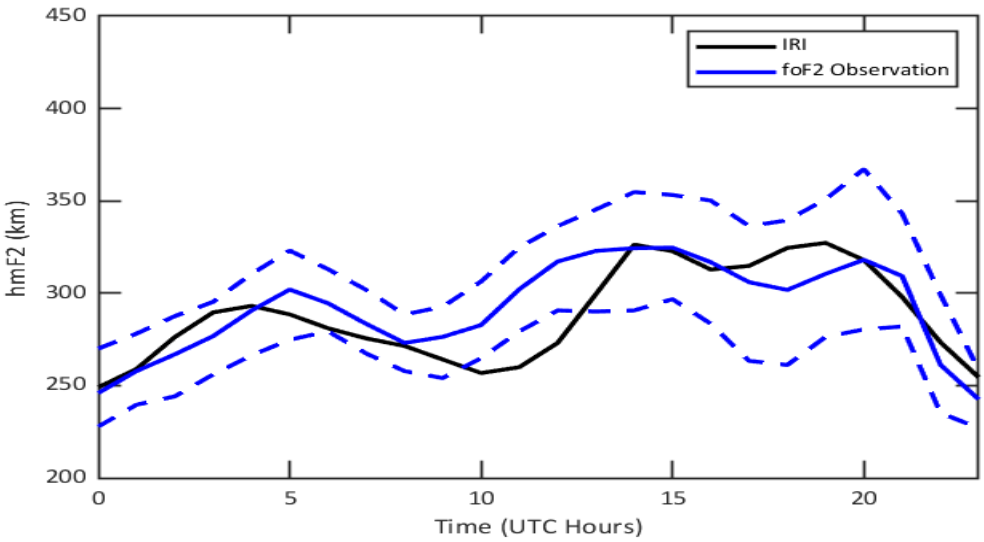

Figure 4.37: A comparison of IRI predictions of hmF2 (black line) with monthly median observations (blue line) and the tenth and ninetieth percentile (dashed blue lines) in March 2004 at Curtin.

4.5 Assessment of IRI's hmF2 model

The performance of IRI's AMTB hmF2 model was analysed. Data from two JORN mid-latitude sites, Curtin (Lat = -18 °N, Long = 124 °E, Dip = -49 °) and Lynd River (Lat = -18 °N, Long = 144 °E, Dip = -47 °), were used. Each site has approximately one solar cycle worth of data. Vanimo and Canberra observations were not used as there are only 18 months of observations available.

An example of a comparison of IRI hmF2 predictions with observations for March 2004 at Curtin is shown in Figure 4.37. This was repeated at Lynd River for Jan 2005 in Figure 4.38. Overall, IRI's hmF2 model appears to correlate fairly well with observations for these two months at these sites except at 10 UT where IRI's prediction was below the tenth percentile.

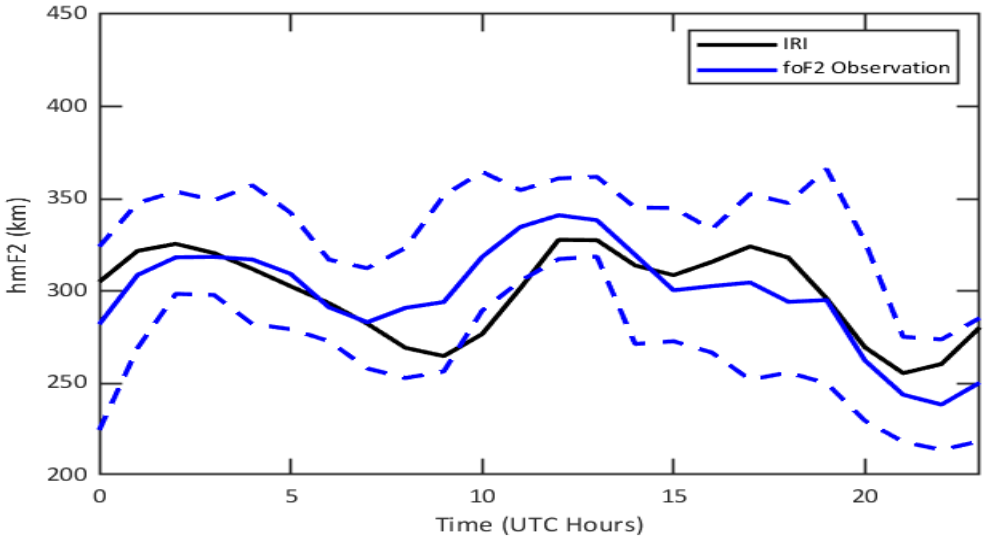

Figure 4.38: A comparison of IRI predictions of hmF2 (black line) with monthly median observations (blue line) and the tenth and ninetieth percentile (dashed blue lines) in January 2005 at Lynd River.

As with the foF2 analysis, the mean and standard deviation of the hmF2 residuals were calculated for each month and are displayed in Figure 4.39. Overall IRI performs well as the standard deviation is consistently below approximately 10 %. During winter solar minimum IRI tends to overestimate hmF2 and during summer solar maximum it tends to underestimate.

This hmF2 residual analysis was repeated at Lynd River, with the results displayed in Figure 4.40. The performance of IRI's hmF2 model is similar at the two sites. The overestimation during winter solar minimum and underestimation during summer solar maximum occurs at both sites.

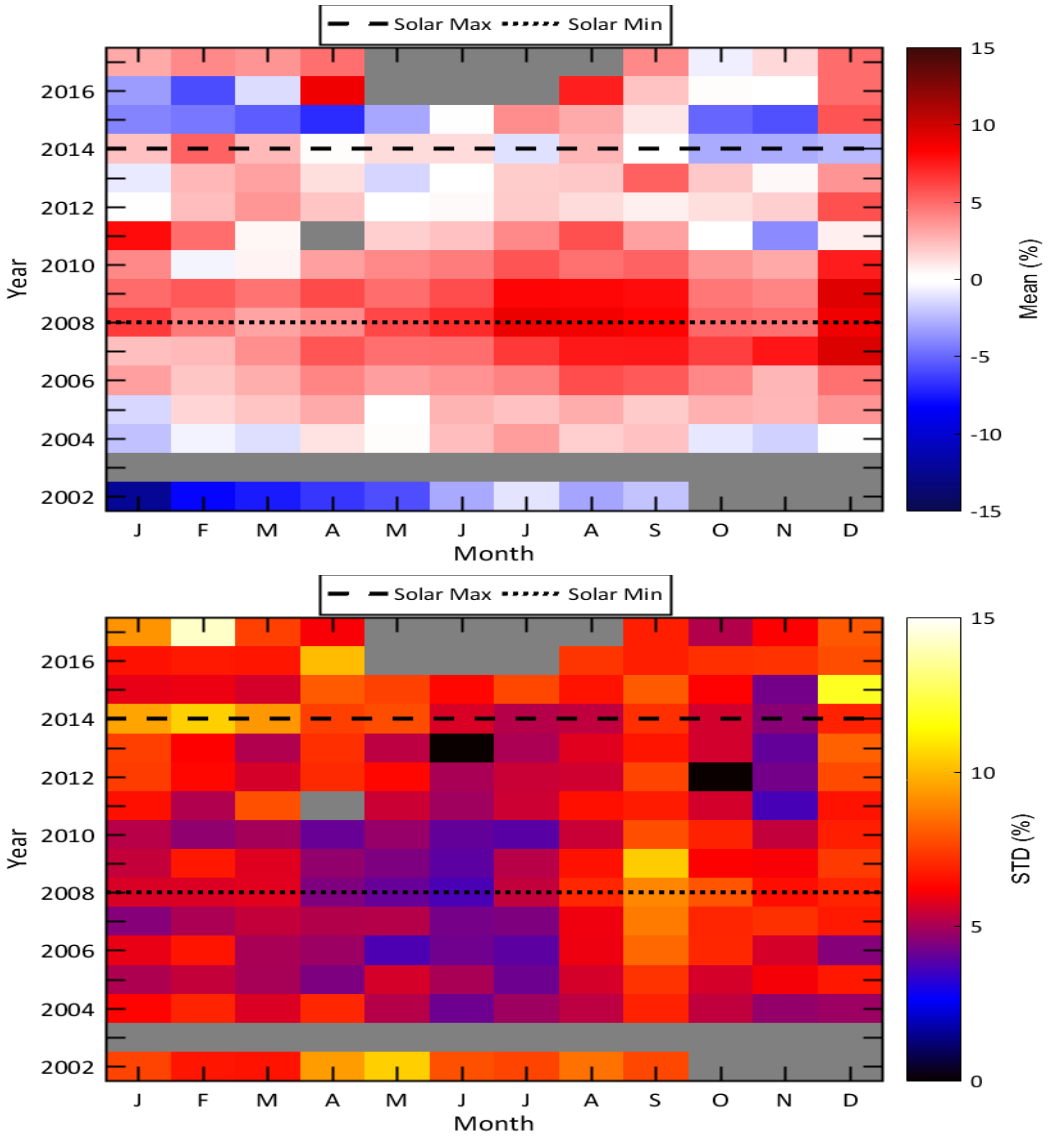

Figure 4.39: The mean (top) and standard deviation (bottom) hmF2 residual from IRI and observations at Curtin. It also indicates solar maximum (long dash line) and solar minimum (short dash line).

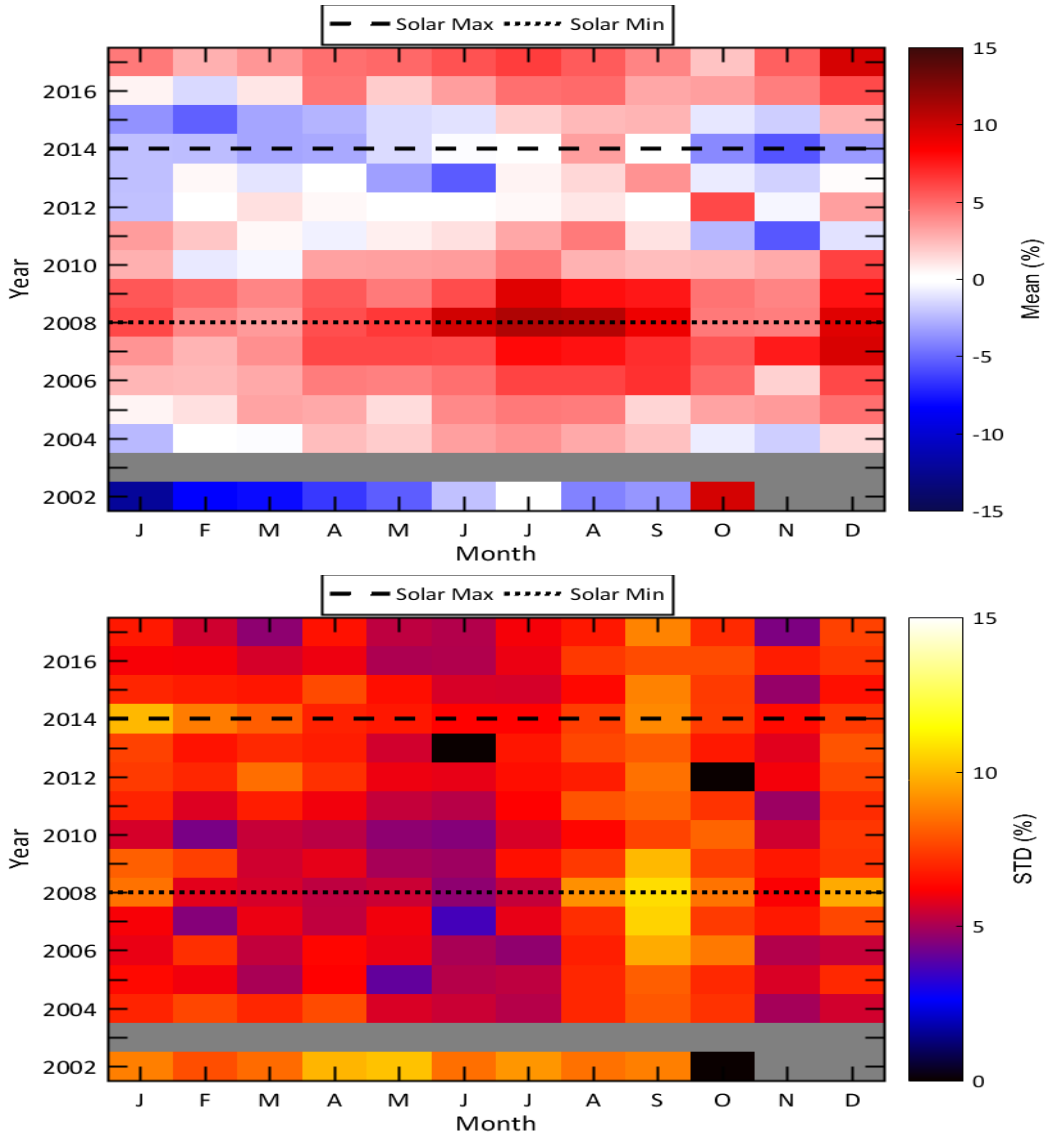

Figure 4.40: The mean (top) and standard deviation (bottom) residual from IRI and hmF2 observations at Lynd River. It also indicates solar maximum (long dash line) and solar minimum (short dash line).

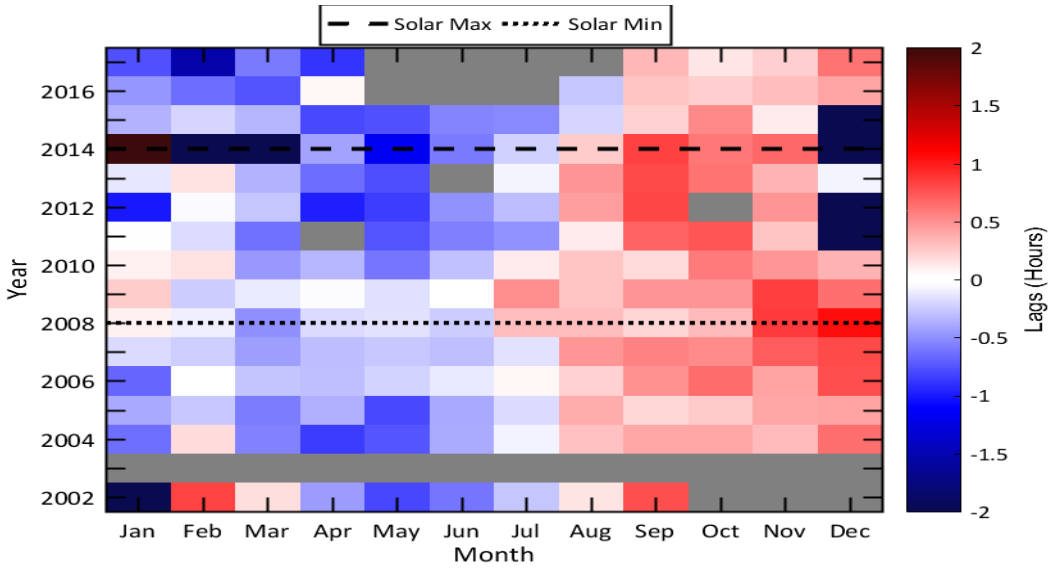

Figure 4.41: Cross correlation lag data gathered from hmF2 observations and IRI at Curtin. It also indicates solar maximum (long dash line) and solar minimum (short dash line).

The cross correlation lag analysis was performed on the hmF2 data from Curtin and Lynd River with results displayed in Figures 4.41 and 4.42 respectively. From these results a seasonal cross correlation lag appears to be present. At both sites, IRI's hmF2 model appears to have a negative lag during the January to July and a positive lag during August to December. There is no apparent solar cycle dependence with these lags.

4.6 Global Analysis of hmF2

Using Juliusruh, Jeju, Puerto Rico, Sanya, Guam, Jicamarca, Scherger, Darwin and Ajana, the global structure of IRI's hmF2 residuals was examined. Figures 4.43 and 4.44 display the mean and standard deviation of the hmF2 residuals. Overall, it appears that IRI overestimates hmF2 at every site, particularly in the Northern Hemisphere, with the most severe overestimation at Sanya and Guam, two northern equatorial anomaly sites. In the Southern Hemisphere it appears that IRI tends to underestimate at solar maximum and overestimate at solar minimum. While there appears to be a seasonal and solar cycle dependence of the mean hmF2 residuals at most of the nine selected sites, it is difficult to discern a latitudinal pattern of behaviour.

Examination of the standard deviation of the hmF2 residual displayed in Figure 4.44 shows the influence of geomagnetic dip. Overall, the standard deviation is fairly low, particularly at Juliusruh, Canberra and Jicamarca. However, the standard deviations are larger around the equatorial anomaly peaks, as shown in Dar-

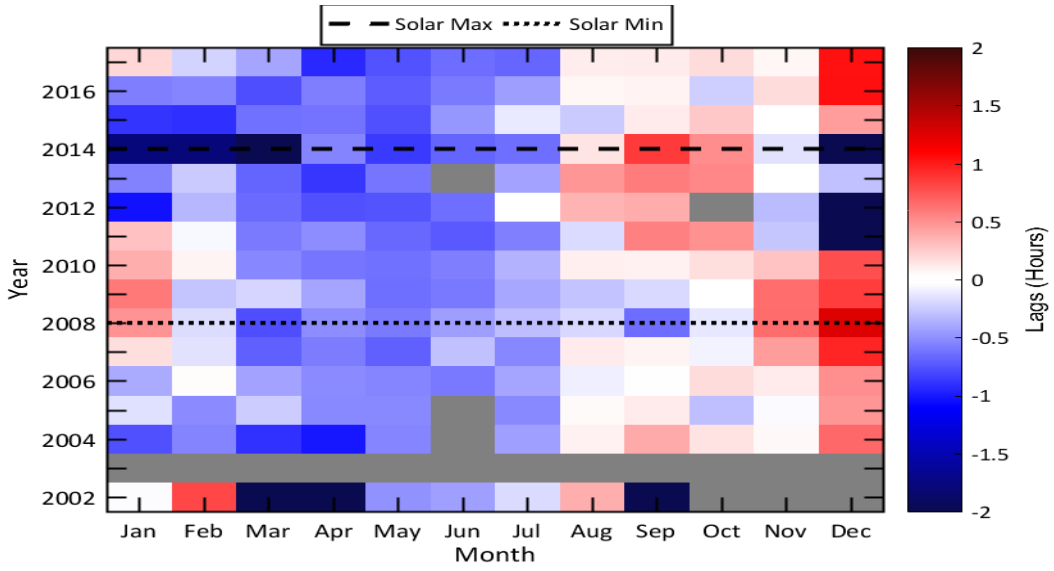

Figure 4.42: Cross correlation lags gathered from hmF2 observations and IRI at Lynd River. It also indicates solar maximum (long dash line) and solar minimum (short dash line).

win, Scherger, Guam, Sanya and Puerto Rico. It appears that IRI performs worst at Darwin, Sanya and Guam particularly during summer, solar maximum. As for Figure 4.43, while there appears to be a seasonal and solar cycle dependence on the standard deviation of the hmF2 residuals it is difficult to discern a latitudinal dependence. Overall, IRI performs better at the mid latitudes and the geomagnetic equator.

The analysis so far has focused on only 9 sites. Figures 4.45 to 4.47 display histograms of the RMS mean and standard deviation for IRI utilising all available sites grouped into different geomagnetic regions. These regions are the Northern Hemisphere mid latitudes (32 contributing sites), low latitudes (23 contributing sites) and Southern Hemisphere mid latitudes (17 contributing sites). IRI's hmF2 model appears to do worse in the low latitudes compared to the mid latitudes. While it appears to perform slightly better in the Southern Hemisphere mid latitudes than the Northern Hemisphere mid latitudes.

The cross correlation lags from IRI's hmF2 model for the nine selected sites are displayed in Figure 4.48. The results are often poor, which is due to IRI's hmF2 model not adequately characterising the temporal behaviour of the monthly median hmF2. An example of IRI not adequately characterising the temporal behaviour of the monthly median hmF2 at Guam in December 2015 is displayed in Figure 4.49. This figure shows that IRI predicts a rise in the ionosphere at sunset and fall at sunrise as expected. However, somewhat surprisingly, this is not seen in the observations. This could be a data quality issue but is beyond the scope of this thesis

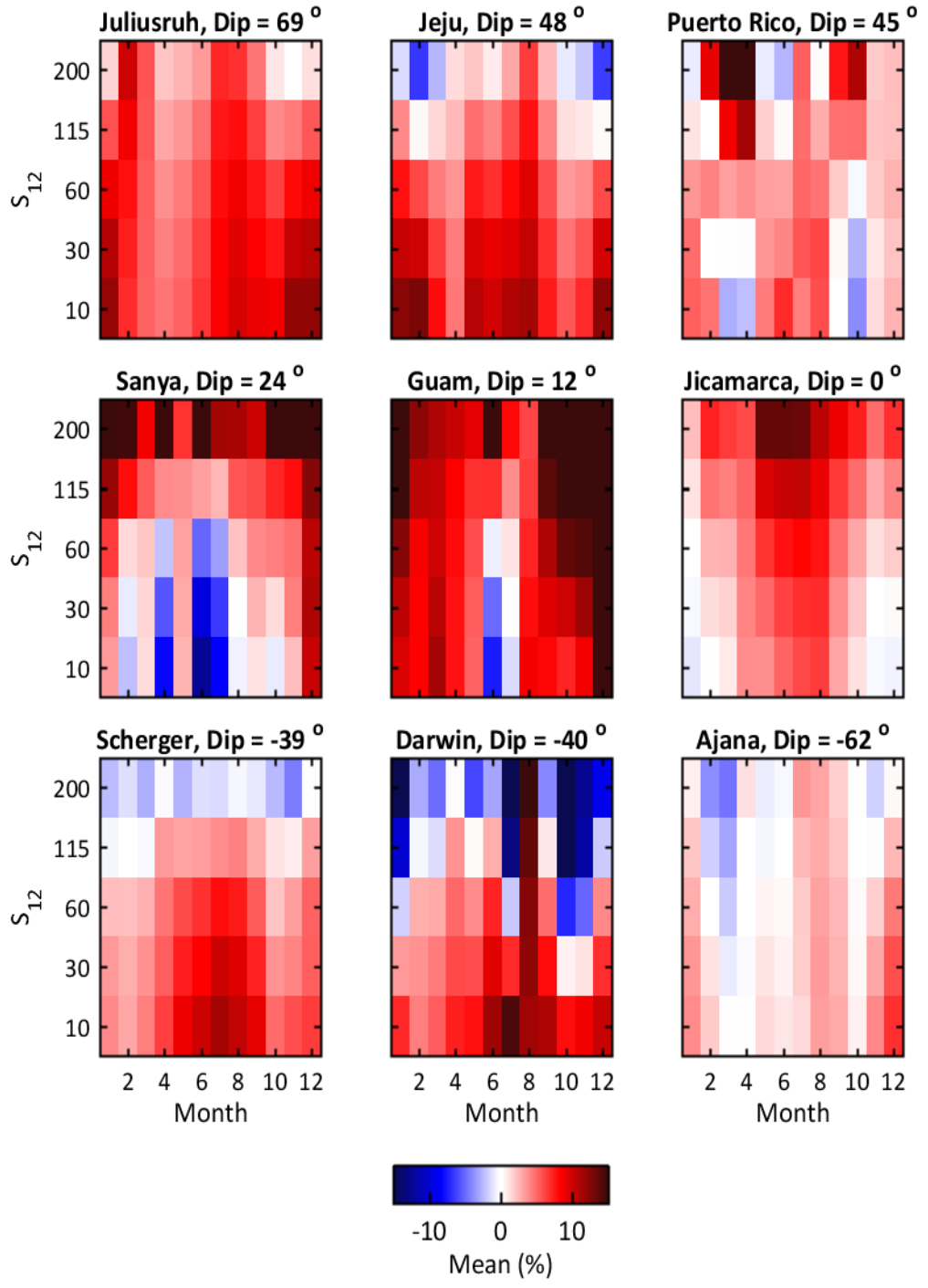

Figure 4.43: The mean hmF2 residual from IRI's hmF2 median model. From left to right, Top: Juliusruh (Dip = 69 °), Jeju (Dip = 48 °), Puerto Rico (Dip = 46 °), Middle: Sanya (Dip = 24 °), Guam (Dip = 12 °), Jicamarca (Dip = 1 °); Bottom: Scherger (Dip = -40 °), Darwin (Dip = -40 °), Ajana (Dip = -63 °) ordered by geomagnetic dip angle.

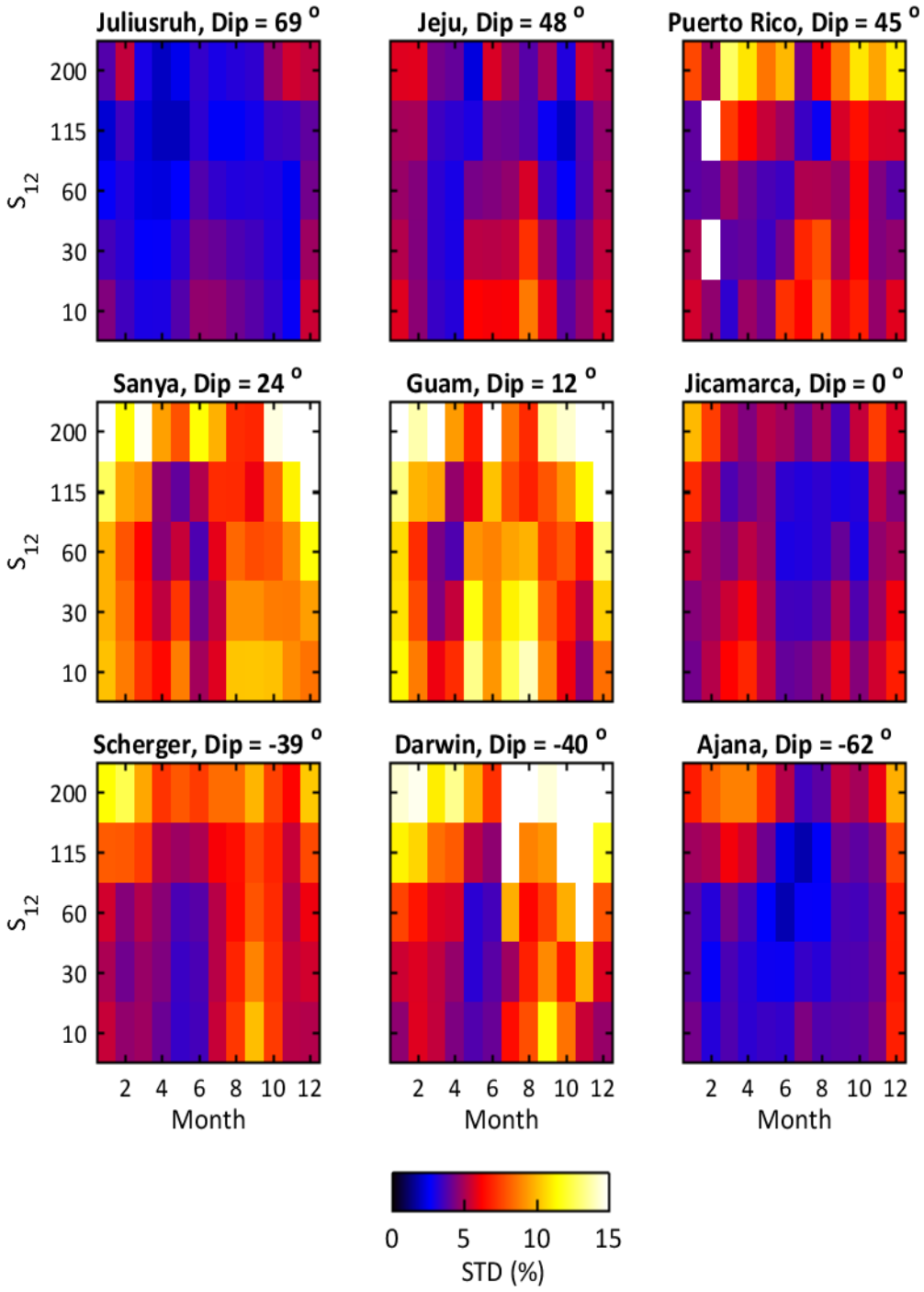

Figure 4.44: The standard deviation of the hmF2 residuals from IRI's hmF2 median model. From left to right, Top: Juliusruh (Dip = 69 °), Jeju (Dip = 48 °), Puerto Rico (Dip = 46 °), Middle: Sanya (Dip = 24 °), Guam (Dip = 12 °), Jicamarca (Dip = 1 °); Bottom: Scherger (Dip = -40 °), Darwin (Dip = -40 °), Ajana (Dip = -63 °) ordered by geomagnetic dip angle.

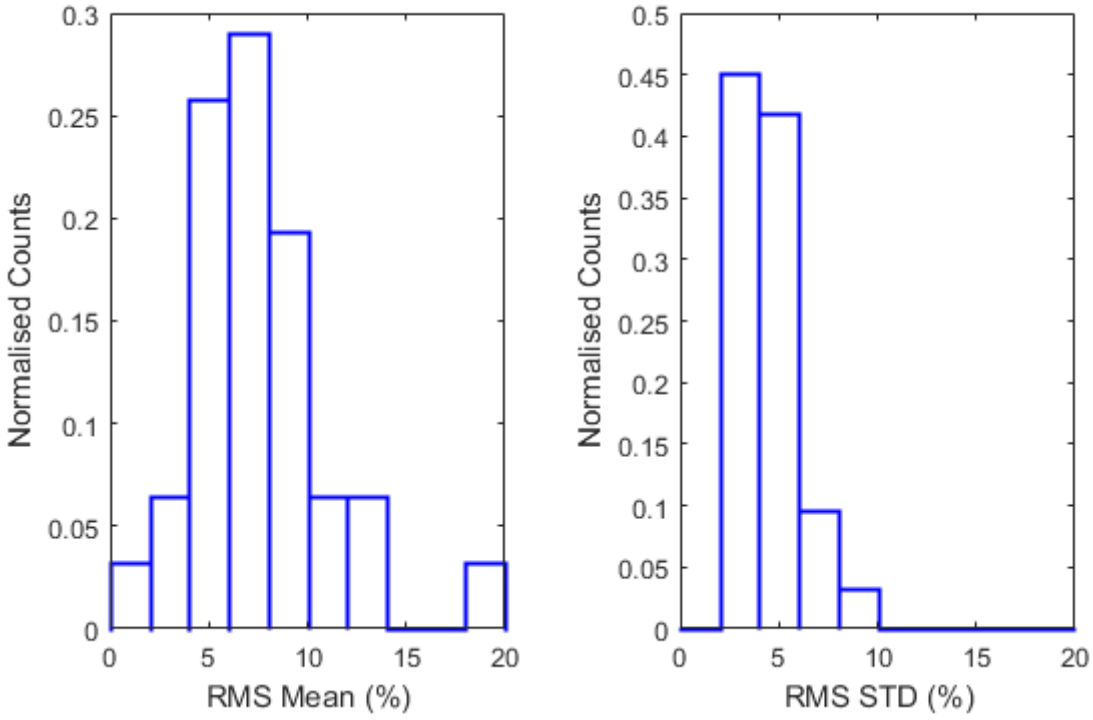

Figure 4.45: RMS of the mean (left) and standard deviation (right) of IRI's hmF2 residuals calculated for the northern mid latitudes (70 ° Dip to 45 ° Dip) using 32 available sites.

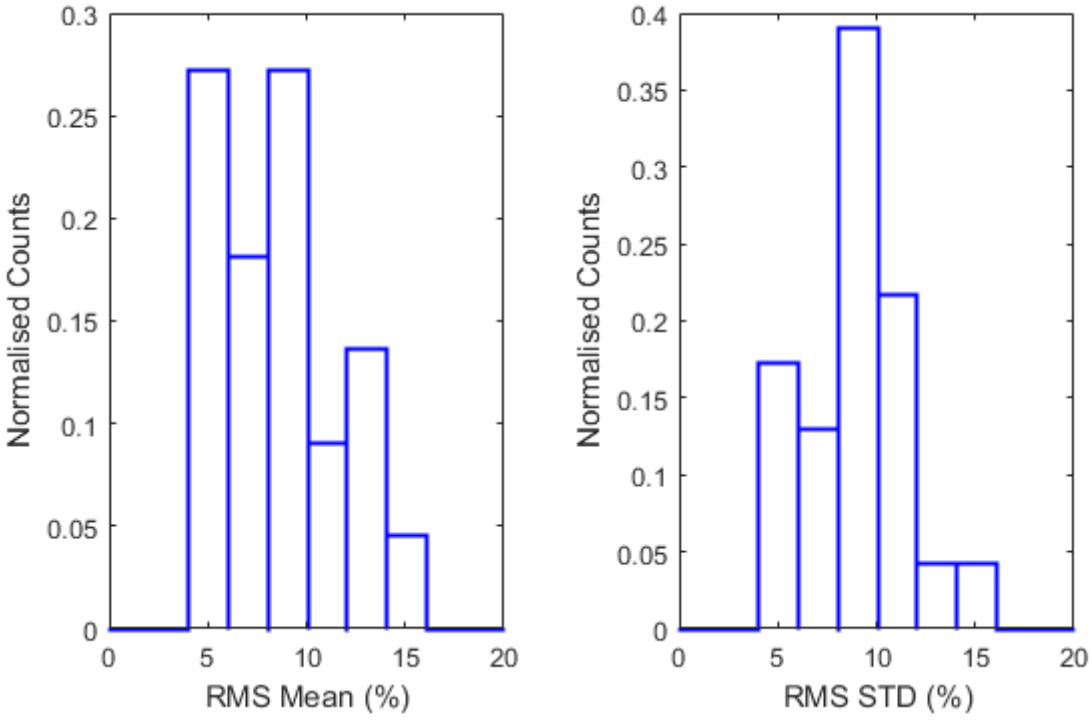

Figure 4.46: RMS of the mean (left) and standard deviation (right) of IRI's hmF2 residuals calculated for the southern mid latitudes (45 ° Dip to -45 ° Dip) using 23 available sites.

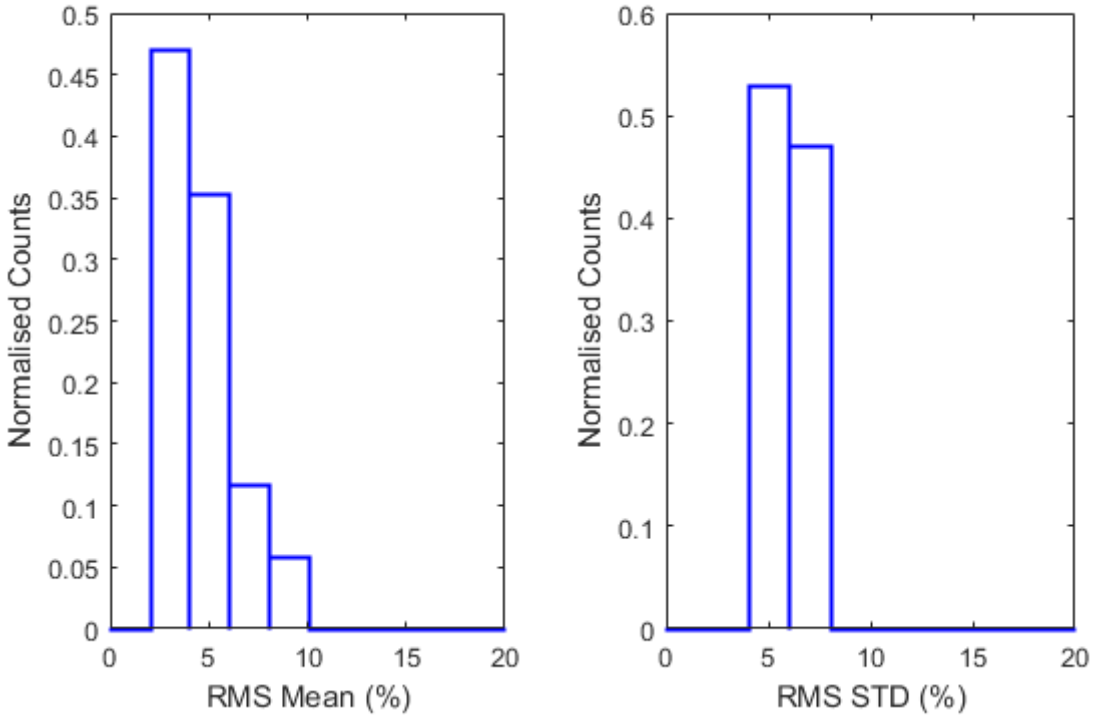

Figure 4.47: RMS of the mean (left) and standard deviation (right) of IRI's hmF2 residuals calculated for the southern mid latitudes (-45 ° Dip to -70 ° Dip) using 17 available sites.

to investigate further. A cross correlation between IRI and the observations for this month produces poor lag results (-11 $\pm$ 0.3 hours). Examination of Juliusruh, Jeju and Puerto Rico indicates that IRI's hmF2 lag is small in the Northern Hemisphere mid latitudes. Further examination of Jicamarca shows a positive lag primarily during solar minimum. At Ajana, there appears to be a small positive lag in winter, solar minimum and a negative lag at solar maximum. Examination of the lag at solar minimum at Darwin and Scherger displays a clear seasonal dependence. Between January to June there is a negative lag, while in July to December there is a positive lag.

4.7 Summary

This chapter examined IRI's default monthly median models of foF2 (URSI) and hmF2 (AMTB). From examination of the residuals between the model and data it is possible to characterise the seasonal, solar cycle and diurnal behaviour of the errors. Overall, it appears that IRI's foF2 model tends to underestimate foF2 during solar minimum and overestimates during solar maximum. Additionally, IRI's foF2 model appears to perform better at the mid latitude sites and at the geomagnetic equator. The largest errors are at the Northern and Southern equatorial anomalies.

There also appears to be a lag in the temporal behaviour IRI's foF2 model, which can be up to 1 hour in magnitude. The lag is worst in the southern hemisphere low latitudes in the Australasian region. The lag is seasonally dependant as it tends to be negative in winter and positive in summer. For the most part there is not a solar cycle influence on the lag. The Northern Hemisphere mid latitude sites tend to have a smaller lag than for SH mid-latitude sites.

The lag in foF2 can be corrected by using a temporal circular rotation. This was done to determine the level of improvement this correction could provide. Overall, this lag correction does not change the mean of the foF2 residuals, but it provides large improvements to the standard deviations of the residuals, particularly at Darwin and Vanimo.

This analysis was also repeated with IRI's hmF2 model. This model tends to overestimate hmF2 at most sites across all seasons and solar cycles, except at solar maximum in the Southern Hemisphere where it tends to underestimate. IRI performs worst at the equatorial anomalies and performs better at mid latitudes and near the geomagnetic equator. The standard deviation of IRI's hmF2 residuals are low at the mid latitudes and geomagnetic equator with standard deviations $\lesssim 5$ %. In the equatorial anomaly regions the standard deviations are ~ 10 %. IRI's hmF2 model appears to have a significant lag problem. The hmF2 lags can vary between -1 to 0.5 hours at sites at Ajana, and -1 to 1 hours at Jicamarca. At some sites (e.g. Sanya) the lag analysis suggests that IRI's characterisation of the diurnal behaviour

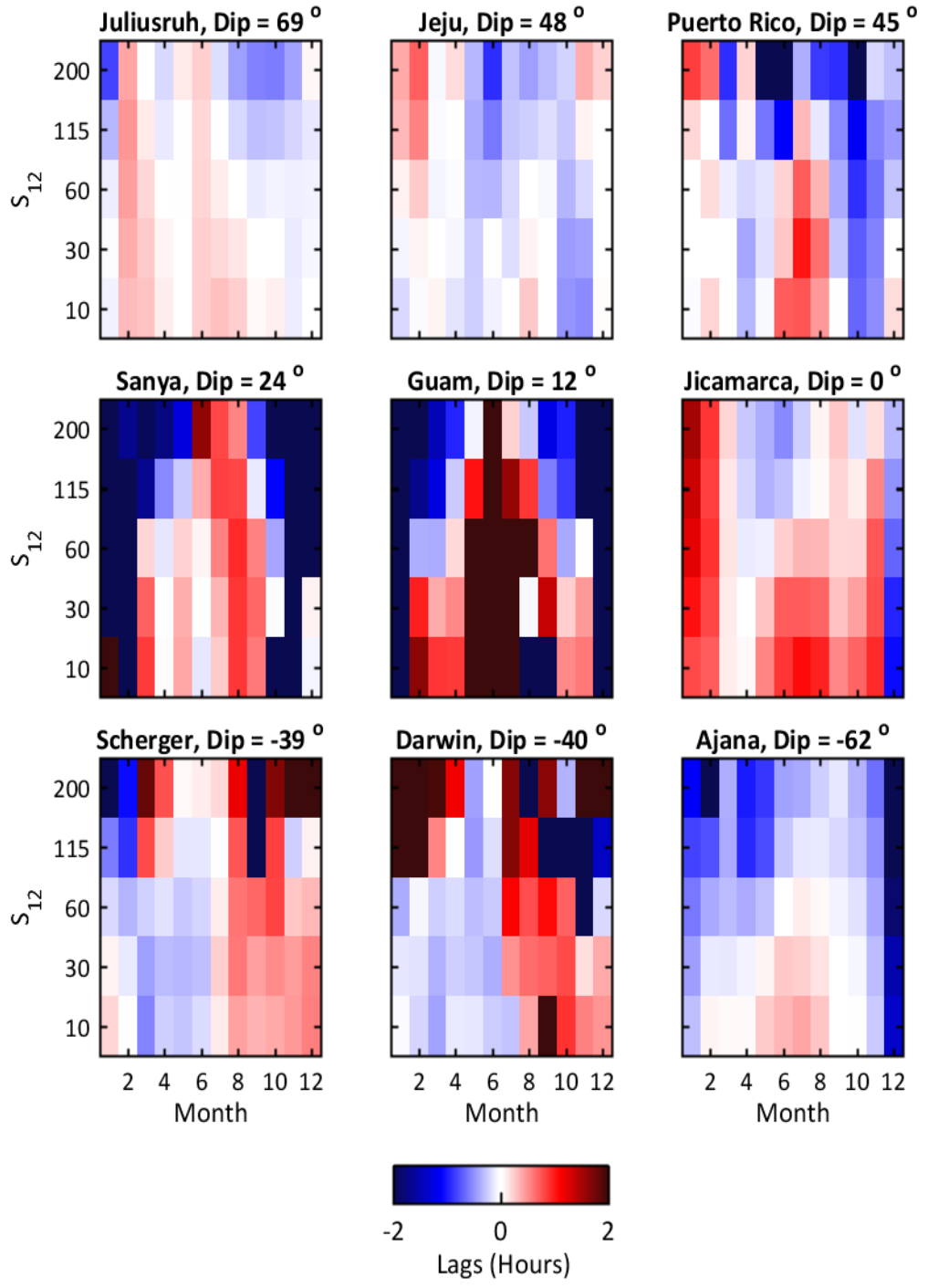

Figure 4.48: The cross correlation lags from IRI's hmF2 median model. From left to right, Top: Juliusruh (Dip = 69 °), Jeju (Dip = 48 °), Puerto Rico (Dip = 46 °), Middle: Sanya (Dip = 24 °), Guam (Dip = 12 °), Jicamarca (Dip = 1 °); Bottom: Scherger (Dip = -40 °), Darwin (Dip = -40 °), Ajana (Dip = -63 °) ordered by geomagnetic dip angle.

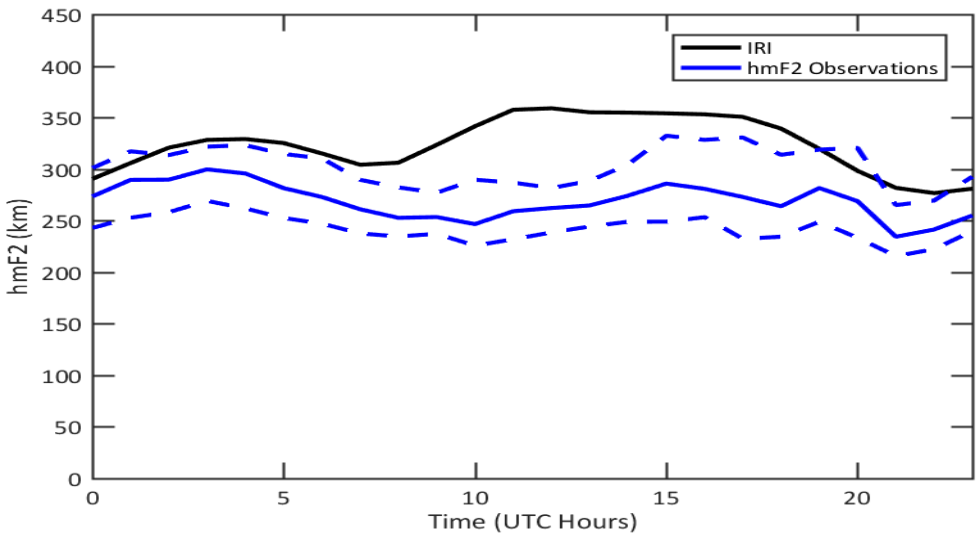

Figure 4.49: IRIs hmF2 model (black line) and monthly median hmF2 (blue line) observations along with the tenth and ninetieth percentiles (dashed blue lines) at Guam in December 2016. Sunrise occurs at 20:44 UTC and sunset occurs at 8 UTC.

of hmF2 is poor.

Chapter 5

A new global climatological model of foF2 and hmF2

The primary shortcomings in the IRI foF2 median model are the lags, tendency to overpredict the strength of the winter southern equatorial anomaly, and the lack of uncertainty estimates. IRIs hmF2 model also has a few shortcomings. It tends to overpredict the height of the F2 layer, it tends to have larger errors in the low latitudes and the lack of uncertainty predictions. To address these shortcomings in IRI a new foF2 and hmF2 model was developed. Note this new model is not intended to replace IRI but to be used in conjunction with it. Three different methodologies were tested to produce three models: Decaf, Latte and Mocha. While the first two models are aimed at providing a regional model in the Australasian region, Mocha provides global predictions of the monthly median foF2 and hmF2 along with their variances.

5.1 Decaf

Decaf produces monthly median, tenth and ninetieth percentile predictions of foF2, using a geographic location, hour, month and year. This model was developed from Australian data.

1. Decaf initially uses the target month and year to find the target F10.7 value from a database.

2. Decaf then finds the closest sounder to the input geographic coordinates and checks if there are at least three monthly median observations for the target month with $0.8(\text{Target F10.7}) \leq \text{F10.7} \leq 1.2(\text{Target F10.7})$.

3. The medians are corrected to the local solar time (LT) of the geographic coordinates.

4. If there are enough data available Decaf takes the median of these observations to predict the target foF2.

5. If insufficient data is available, it finds the next closest sounder and repeats the above process. When searching for sounders it uses an elliptical search area, increasing the search radius by 2 degrees longitude and 1 degree latitude.

Decaf has several limitations that make it less preferable to IRI.

1. Decaf cannot produce predictions if there are no nearby sounders (within 30 ° Lat and 30 ° Long) and thus is unusable in oceanic regions or other regions lacking sounders.

2. Decaf also lacked a latitude correction to correct for different times of day, which caused severe under or over estimation problems.

5.2 Latte

Latte was developed to remediate Decaf's flaws by using a different methodology. Instead of relying on a single sounder, Latte fits polynomials to data from a collection of sounders.

Latte's fit functions require representing the sounder sites in terms of either geographic latitude, magnetic dip angle or modified dip angle. Representing the sites by geographic latitude would not represent the magnetic field influence and thus would not represent the ionosphere accurately. The Earth's magnetic equator varies in geographic latitude across the globe. A map of the geomagnetic dip equator and $\pm 30\,^\circ$ dip angles is displayed in Figure 5.1. From Africa to Indonesia, the magnetic equator is held at an almost fixed geographic longitude of $\sim 8\,^\circ$ N. However, between the eastern side of Australia and South Africa, it slowly dips southwards, before sharply rising Northward towards Africa. Sorting sites by dip or modified dip angle would provide a more accurate representation of the ionosphere than ordering by geographic latitude. As Bilitza (1990, p. 52) explains, modified dip angle provides a more accurate representation of the ionosphere than dip.

A variety of fitting types such as polynomial, Legendre, or Chebychev polynomials were investigated. According to Jones and Gallet (1962), polynomials and Legendre polynomials diverge quickly at their endpoints, however, Chebychev polynomials aim to minimise the maximum errors, and therefore perform better at the endpoints. Therefore, Chebychev fits were chosen for this analysis, but as displayed in Figure 5.2, it is clear that different fit types provide similar results.

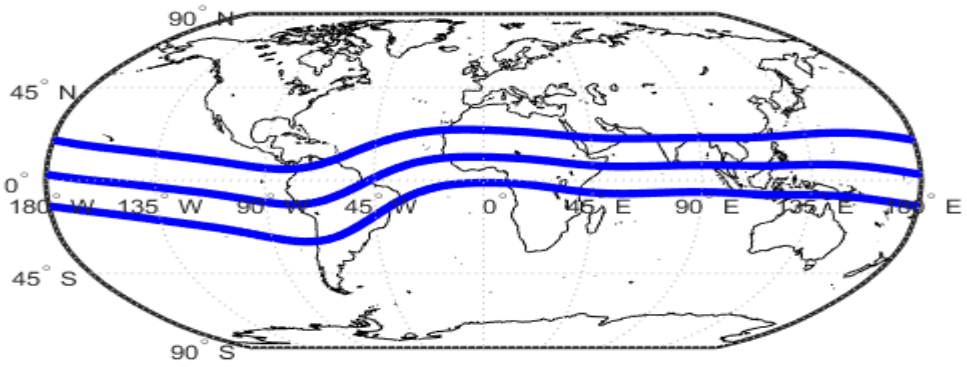

Figure 5.1: Blue lines indicate the geomagnetic dip equator and ± 30 ° geomagnetic dip angles.

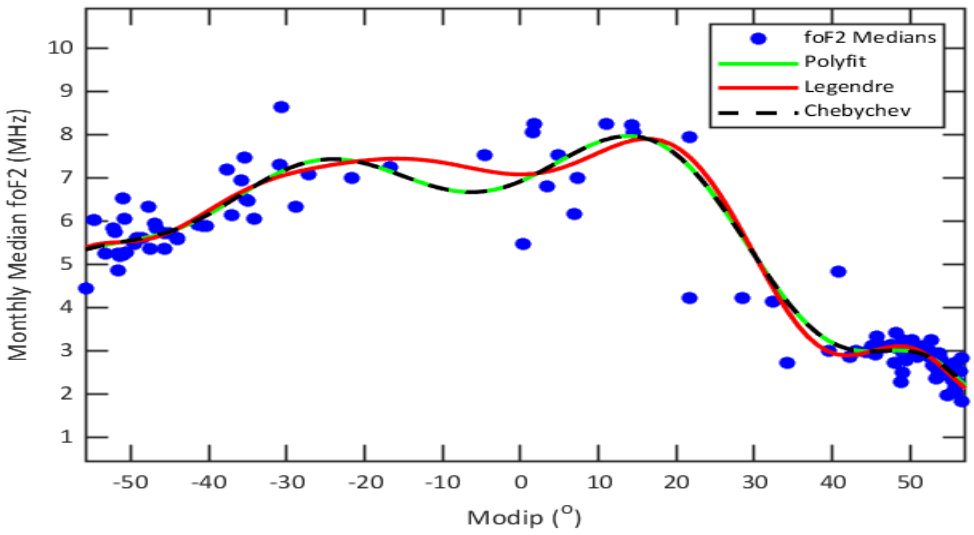

Figure 5.2: Example of different fits applied to foF2 median data from January in the S_{12} bin $= [0, 20]$ at 19 LT. The fits are a 12th order Legendre (red line), a 10th order Chebychev (dashed black line) and a 10th order polynomial fit (green line).

Finally, the order of the Chebychev fit was chosen using MATLAB's goodness of fit test (*Mathworks*, 2016). A fit was attempted at every order from 1st order to 14th order, MATLAB then evaluates the goodness of fit at each order using the root mean square error, RMSE, defined as

$$\text{RMSE} = \sqrt{\frac{SSE}{n-m}}, \tag{5.1}$$

where n is the number of response values, m is the number of fitted coefficients and SSE is the sum of the squared residual.

After the fitting types and sorting parameters were determined Latte would predict foF2 and hmF2 for a target time and geographic coordinates. Like Decaf, Latte was limited to the Australian region. This was done as follows:

1. Latte has access to a database of local time (LT) corrected monthly median, tenth and ninetieth percentiles foF2 observations at a large number of sites. This local time correction was done to provide a longitudinal correction to the foF2 and hmF2 fitting. This correction was discussed in more detail in Section 4.2.

2. Firstly, Latte sorts the sites by modified dip angle to allow accurate description of the geomagnetic latitude variation of foF2.

3. Secondly, it performs a third order Chebychev fit to the median, tenth and ninetieth percentiles foF2 observations over modified geomagnetic dip angle for each hour in the monthly median, tenth and ninetieth percentiles, between 1980 to 2014 to provide a modified dip correction. A third order fit was chosen to avoid overfitting as the data were typically limited (< 13 foF2 observations).

4. To determine the best fit to use, Latte converts the target geographic coordinates to a target modified dip angle and an F10.7 value.

5. Finally, Latte finds the year with closest F10.7 to the target F10.7, and evaluates the polynomial for this time at the target modified dip angle to produce the predicted foF2.

Unfortunately Latte's methodology has a few difficulties:

1. If the requested year/month combination lacked data or the third order polynomial fit was poor then it would produce inaccurate predictions. To address this problem, Latte would search for the next best year, which may have data but with a potentially worse fit.

2. Lower latitudes were problematic, as there was insufficient data to describe this region.

Sort Type	Mean foF2 Residual (%)	STD foF2 Residual (%)
modified dip angle	1.6	11.0
Dip	1.8	11.3

Table 5.1: An example of each sorting type on foF2 site data using a Chebychev fit, tested at Jicamarca, including Jicamarca in the model. The mean and standard deviation of the foF2 residual are calculated over the entire month range, over the entire solar cycle.

3. The fits would rapidly diverge from the actual values if the target coordinates were in the Northern Hemisphere as Latte is based on Australian data.

4. Due to the limited number of data it was possible for the tenth (or ninetieth) percentile to be greater (smaller) than the fiftieth percentile, which is not physical.

5.3 Mocha

Mocha expands on Latte's methodology for a more accurate global climatological model. While Latte was limited to the Australia region by filtering sites through their geographic coordinates, this restriction was removed in Mocha. This allowed for higher accuracy in low latitudes due to the increase in the number of sites. Additionally, it could extend accurately into the Northern Hemisphere.

Mocha's global fits were tested using both dip and modified dip angle sorting. Table 5.1 presents the residuals obtained on applying a Chebychev fit across magnetic dip angle and modified dip angle. Examination of Mocha's performance at Jicamarca indicates that modified dip angle is the better option, with lower residuals.

Mocha's methodology is summarised as;

1. For each sounder, month and hour, the local time corrected foF2 and hmF2 data is binned into solar regimes, S_{12} = [0, 20], [20, 40], [40, 80], [80, 150] and [150, 250]. These bins were chosen based on data availability across all sunspot regimes, as less data was available at solar maximum, necessitating the use of larger bin sizes for larger S_{12}.

2. Linear interpolation is used to correct the foF2 and hmF2 observations to the centre of each S_{12} bin, thus providing foF2 estimates at S_{12} = 10, 30, 60, 115 and 200.

3. For each month, hour and S_{12} value, a Chebychev polynomial fit is applied to local time corrected observations across modified dip angle as in Section 4.2.

4. The target geographic coordinates is converted to a target modified dip angle and the target year converted to the target S_{12}.

5. The S_{12} fits are solved for S_{12} = 10, 30, 60, 115 and 200 at the target time and modified dip angle.

6. Linear interpolation is performed across S_{12} to solve for any target S_{12}.

7. Linear interpolation between months is performed to get foF2 predictions at any day.

Using S_{12} or F10.7 as a measure of solar activity was undertaken during the development of the Latte model with no significant advantage of one parameter over the other being discernible. The choice of S_{12} for Mocha based upon experiences with Latte was deemed to be arbitrary. Further investigation of using F10.7 in place of S_{12} within the Mocha model could provide additional insight but was beyond the scope of this thesis.

Mocha's methodology attempts to address the limitations found in Latte. Latte has difficulties due to the inconsistency in the data fitting as some hours have little data available. By binning the data by sunspot number, more foF2 data is available, allowing higher order polynomial fits to be applied.

5.4 Mocha Analysis At An Independent Site

Mocha was tested at Darwin, a mid to low latitude site not included within the foF2 fitting but included in IRI's median foF2 model. Note that Darwin was removed from the Mocha fits only for this comparison. Using the same analysis process as discussed in Chapter 4, summary plots of Mocha and IRI's foF2 mean residual, standard deviation and lag are given in Figures 5.3, 5.4 and 5.5 for Darwin.

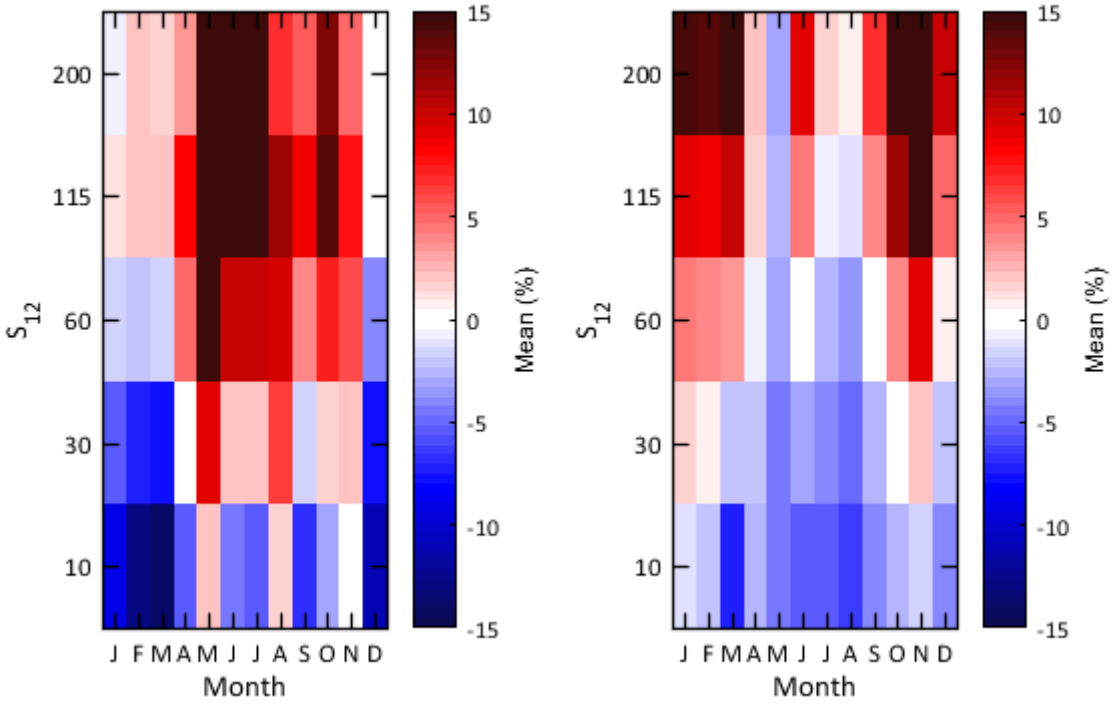

Figure 5.3: A comparison between the mean foF2 residual at Darwin as a function of solar activity and month for IRI (left) and Mocha (right).

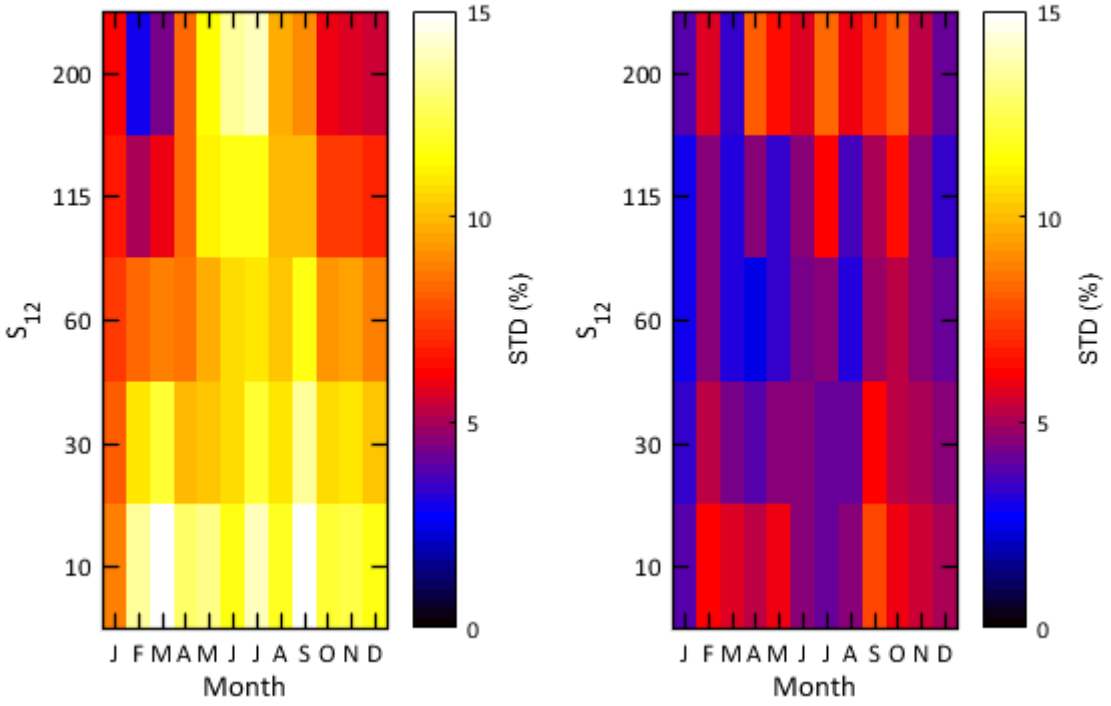

Figure 5.4: A comparison between the standard deviation of the foF2 residual at Darwin as a function of solar activity and month for IRI (left) and Mocha (right).

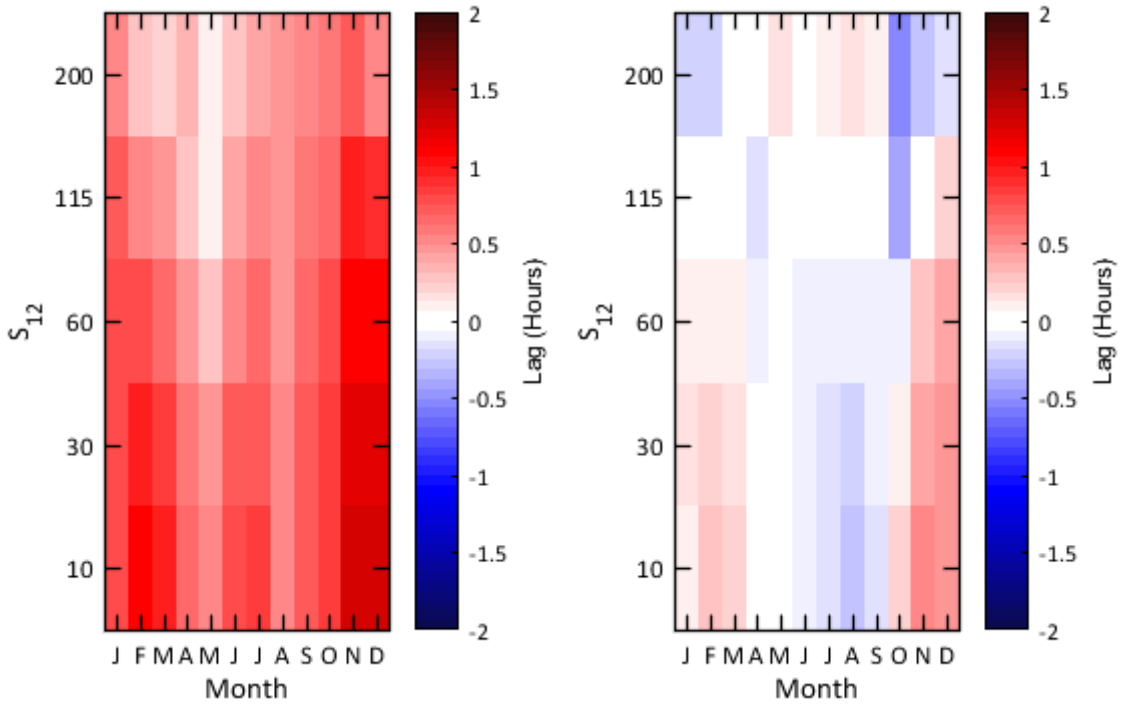

Figure 5.5: A comparison between the foF2 cross correlations lags at Darwin as a function of solar activity and month for IRI (left) and Mocha (right).

Comparing the performance of the two models shows that Mocha's median foF2 model outperforms IRI's median foF2 model at Darwin with lower residuals. From the mean residual results in Figure 5.3, it can be seen that Mocha works better at low sunspot numbers than high sunspot numbers. At low sunspot numbers Mocha performs slightly better in November to March (i.e. Southern Hemisphere summer) than May to August (i.e. Southern Hemisphere winter) and during high sunspot numbers it performs better in (Southern Hemisphere) winter than summer. IRI performs better than Mocha at solar maximum during November to March.

Examination of the standard deviation of the foF2 residuals in Figure 5.4 shows similar improvements. Mocha outperforms IRI in every month and sunspot number except for February, $S_{12} = 200$. It appears that Mocha performs best for medium sunspot numbers ($S_{12} = 60$) than at high or low sunspot numbers. It appears to performs worst at solar maximum ($S_{12} = 200$).

Finally, examination of the foF2 lags in Figure 5.5 shows that Mocha has a smaller lag than IRI. The lag in Mocha is only slightly positive during November to December and slightly negative in October. In comparison, the lag in IRI tends to be positive throughout the entire solar cycle.

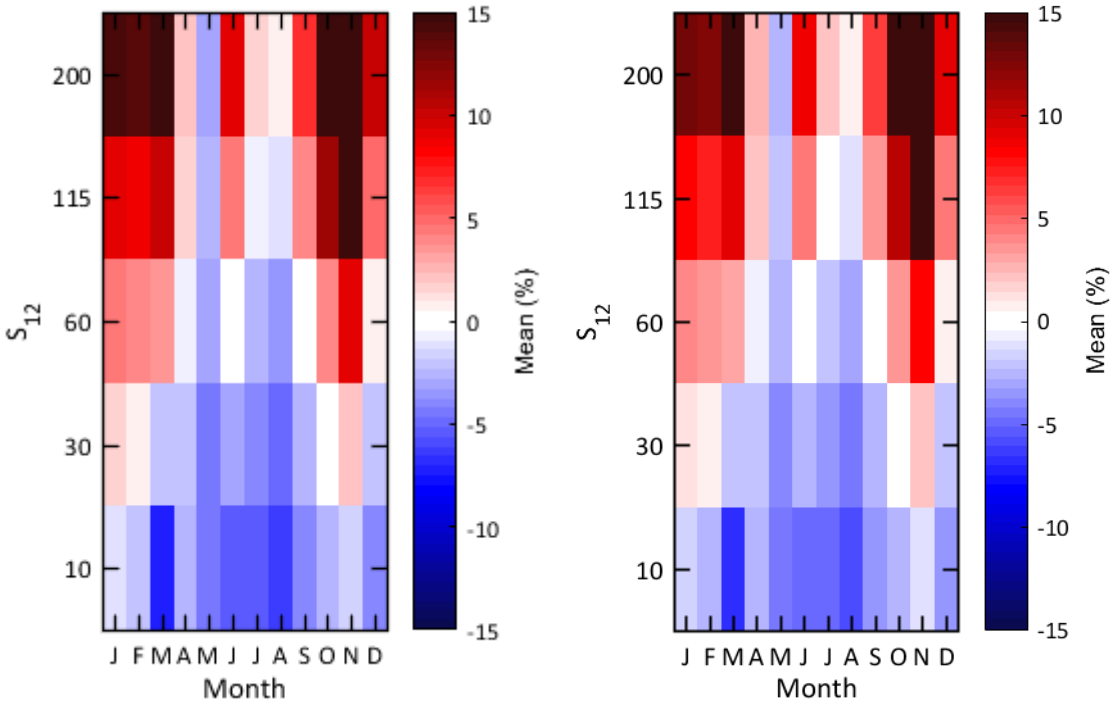

Figure 5.6: A comparison between the mean foF2 residual at Darwin for Mocha without including Darwin in the foF2 median fits (left) and Mocha including Darwin in the fits (right).

Overall, Mocha's foF2 model outperforms IRI's foF2 model at Darwin by a considerable margin, with significant improvements to the standard deviation, but there are only slight improvements to the accuracy. Additionally, Mocha also does not appear to have a lag problem like IRI does.

5.5 Mocha Analysis Including An Independent Site

To determine the effect of including Darwin foF2 observations in the foF2 median fits, Mocha was reapplied incorporating Darwin foF2 observations. Figures 5.6, 5.7 and 5.8 display the recalculated mean and standard deviation of the foF2 residuals and lag results, indicating that including Darwin observations only provides a slight improvement. The remainder of this analysis includes all sites within Mocha's tenth, fiftieth and ninetieth percentile foF2 fits to improve accuracy.

5.6 Mocha and IRI Comparisons

Table 5.2 displays the RMS of the mean and standard deviation of the Mocha and IRI residuals. Examination of this table indicates the improvements between IRI's median foF2 model and Mocha's median foF2 model. Mocha has a lower RMS

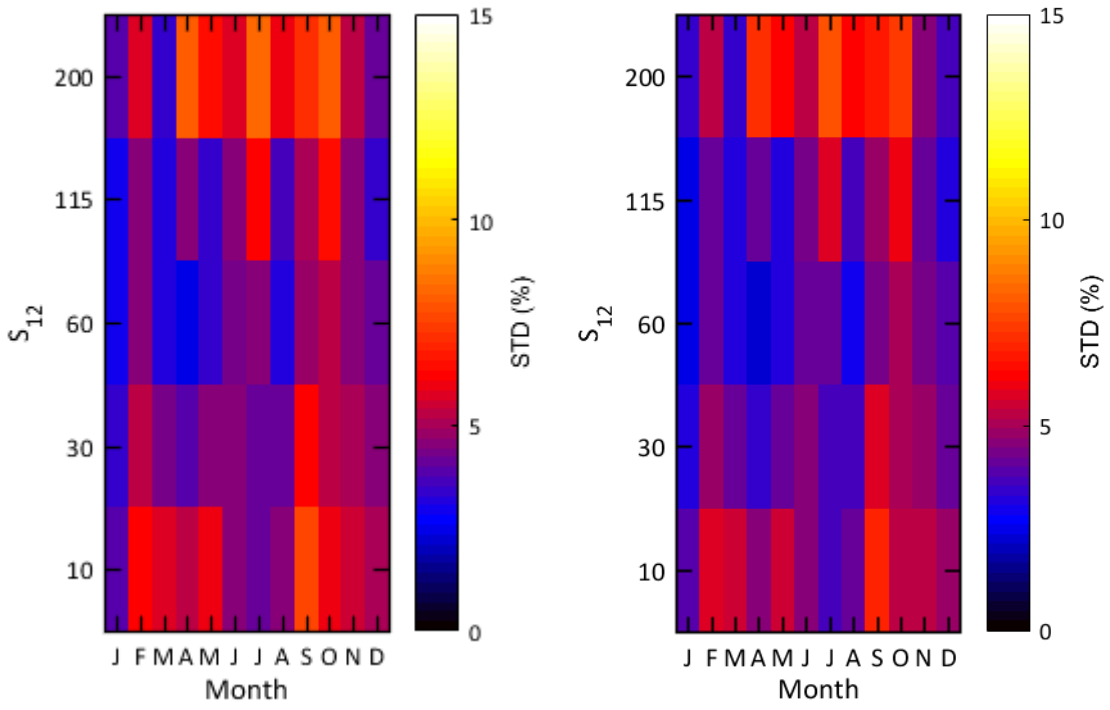

Figure 5.7: A comparison between the standard deviation of the foF2 residual at Darwin for Mocha without including Darwin in the foF2 median fits (left) and Mocha including Darwin in the fits (right).

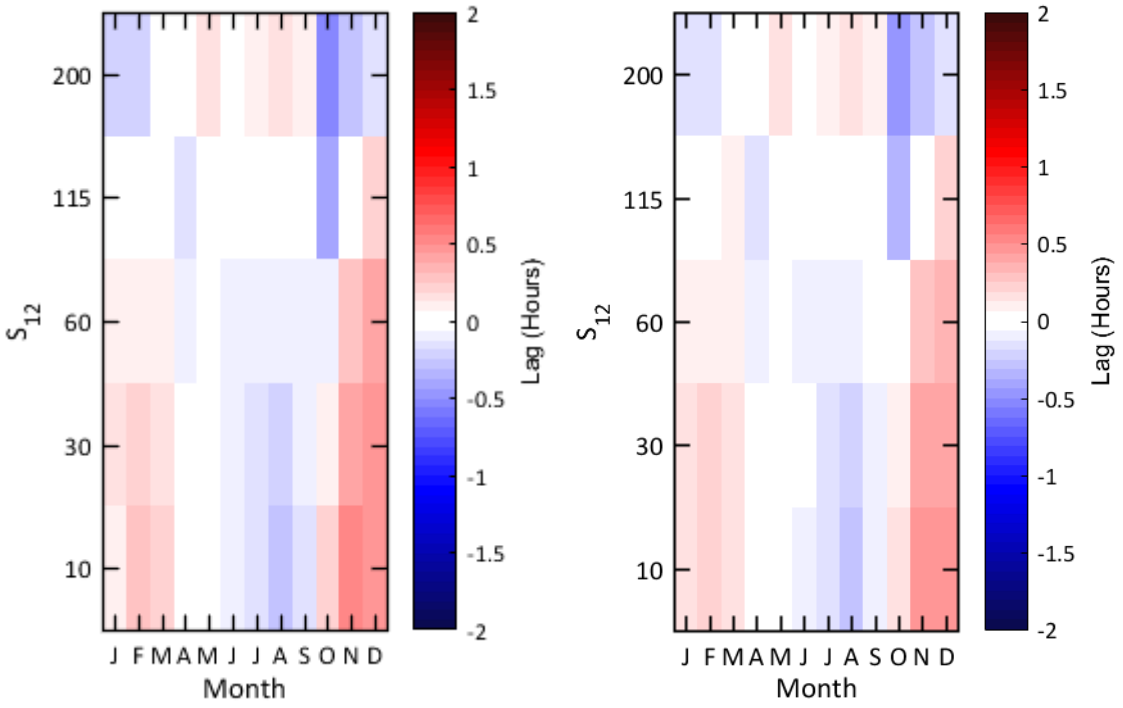

Figure 5.8: A comparison between the foF2 lag at Darwin for Mocha without including Darwin in the foF2 median fits (left) and Mocha including Darwin in the fits (right).

mean residual than IRI at 6 sites, Darwin, Port Moresby, Jicamarca, Puerto Rico, Tokyo and Juliusruh, while IRI had a lower RMS mean at Canberra, Vanimo and Djibouti. Mocha had a lower RMS standard deviation than IRI at 7 sites, Darwin, Port Moresby, Vanimo, Jicamarca, Djibouti, Tokyo and Juliusruh, while IRI had a lower RMS standard deviation at Canberra and Puerto Rico. It appears the Mocha's median foF2 model is slightly more accurate than IRI's median foF2 model and so it is an improvement. It appears that the improvements to the RMS mean and standard deviation of the residuals in Mocha over IRI occurs at low latitude sites, where IRI had the greatest difficulties, as shown in Chapter 4.

The solar cycle summary plots for Mocha's foF2 median model for the mean and standard deviation of residuals and lag are displayed in Figures 5.9, 5.10 and 5.11 respectively. For a comparison, IRI's foF2 model mean, standard deviation and cross correlation lags are displayed in Figures 4.14, 4.15 and 4.27.

Examination of the mean residuals from Mocha's median (fifty percentile) foF2 model in Figure 5.9 shows the influence of the season, solar cycle and geomagnetic dip on the residual for some sites. At Juliusruh, Tokyo, Djibouti, Port Moresby and Canberra there does not appear to be a significant solar cycle influence. At Puerto Rico, Jicamarca and Darwin, it appears that Mocha's foF2 model underestimates (negative mean) during solar minimum and overestimates (positive mean) at solar

Site	Dip (°)	IRI RMS Mean Residual (%)	Mocha RMS Mean residuals (%)	IRI RMS STD of residuals (%)	Mocha RMS STD of residuals (%)
Canberra	-66	5	5.6	6	7.9
Darwin	-40	8.9	6.7	10.3	4.7
Port Moresby	-33	9.2	5.9	9.7	7.1
Vanimo	-22	4.9	6.2	9.7	6.6
Jicamarca	1	6.4	6.2	10.2	9.2
Djibouti	11	5.6	7.6	10.6	9.8
Puerto Rico	46	12.3	10.6	8	9.5
Tokyo	49	7.7	4.1	7.8	4.9
Juliusruh	69	6.5	2	5.6	3.3

Table 5.2: A comparison between IRI's foF2 model and Mocha's foF2 model at a few selected sites.

maximum, and by a large margin at Puerto Rico compared with other sites. At Vanimo, the reverse occurs, with overestimation at low sunspot numbers and underestimation at high sunspot numbers. It appears that Mocha has a seasonal bias at some sites, with underestimates during winter and overestimates during summer at Canberra, Darwin and Vanimo. Overall, Mocha's foF2 model appears to do worst at the Northern Hemisphere low latitudes as seen by Jicamarca, Djibouti and Puerto Rico, and performs best at the Northern Hemisphere mid latitudes in Juliusruh and Tokyo.

Figure 5.10 shows the seasonal, solar cycle and geomagnetic influence on Mocha's median foF2 model standard deviations of the foF2 residuals. At these specific sites, the residuals are only weakly influenced by the solar cycle, with a slight indication at Puerto Rico, Djibouti, Jicamarca and Canberra. Some sites appear to have a no seasonal dependence, for example Canberra, Darwin, Tokyo and Juliusruh, which are all mid latitude sites. At low latitude sites Port Moresby and Puerto Rico, it appears that Mocha's median foF2 model performs better during each sites local summer than their local winter. Overall, Mocha's foF2 model performs best at the Northern Hemisphere mid latitude sites, Tokyo and Juliusruh, which agrees with the results from the mean residual.

Figure 5.11 displays the seasonal, solar cycle and geomagnetic dip angle influence on Mocha's median foF2 model cross correlation lags. With the exception of Canberra in summer at mid to high solar activity, the cross correlation lags are less than 1 hour at every site. Some sites tend to have a predominately positive lag, like Puerto Rico, and others sites like Port Moresby have a predominately negative lag and there is not a distinct hemisphere bias. With the exception of Canberra, there is not a large solar cycle or seasonal influence on the lags, and Canberra has

a seasonal effect (summer lag) that appears to be enhanced by the solar cycle, such that increasing ionising flux represented by S_{12} always increases the positive lag.

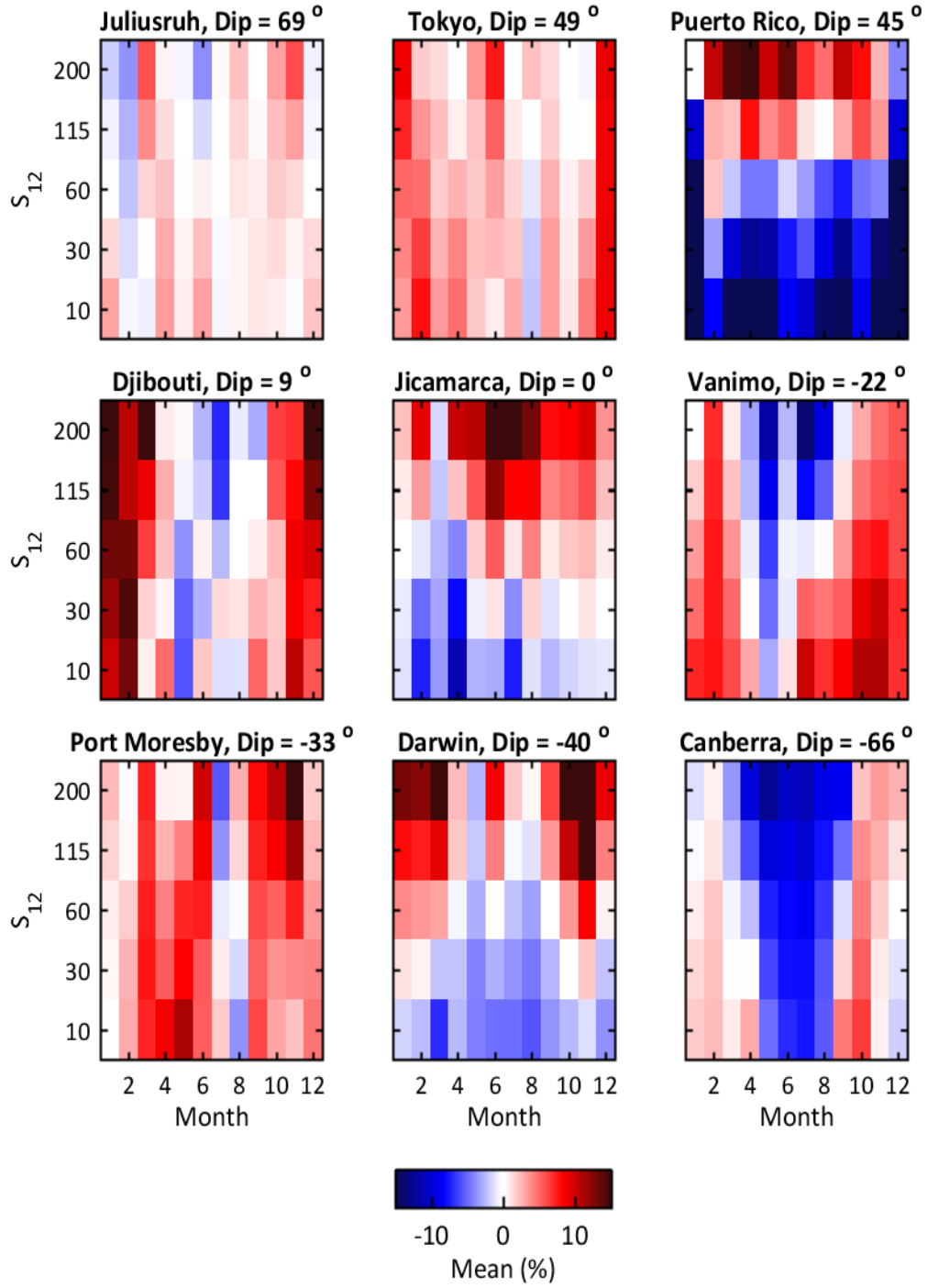

Figure 5.9: The mean foF2 residual between Mocha's fifty percentile foF2 model and monthly median observations for several sites ordered by geomagnetic dip angle. From left to right, Top: Juliusruh (Dip = 69 °), Tokyo (Dip = 49 °), Puerto Rico (Dip = 46 °); middle: Djibouti (Dip = 11 °), Jicamarca (Dip = 1 °), Vanimo (Dip = -22 °); bottom: Port Moresby (Dip = -33 °), Darwin (Dip = -40 °), Canberra (Dip = -66 °)s.

Figures 5.9 and 4.14 provide a comparison between Mocha's foF2 median model residual mean and IRI's foF2 median model residual mean. Both models perform differently enough to discriminate between them at every site, with the exception of Jicamarca and Puerto Rico, where they have similar solar minimum underestimation and solar maximum overestimation, with little seasonal influence. At Juliusruh, Tokyo and Port Moresby, Mocha's foF2 model has improved performance over IRI at all times. At Darwin, Mocha performs better during solar minimum and winter solar maximum but IRI performs better during summer at solar maximum. At Canberra, Mocha performs better during summer, solar minimum, and IRI performs better at winter solar maximum. They perform similarly at other times. At Vanimo, IRI has an improved mean residual at all times except during winter, solar minimum. Finally, at Djibouti, Mocha outperforms IRI during May to October, but IRI performs significantly better during November to February.

Examination of Figures 4.15 and 5.10 shows the comparison of IRI's and Mocha's median foF2 model standard deviation of residuals at a few selected sites. Mocha's foF2 model always provides an improvement over IRI's foF2 model at Juliusruh, Tokyo, Darwin, Port Moresby and Vanimo. In general Mocha's median foF2 model provides an improvement over IRI at Jicamarca and Djibouti except for during solar maximum. IRI tends to perform better than IRI except during summer solar maximum at Puerto Rico.

Figures 4.27 and 5.11 provide a comparison between IRI's and Mocha's median foF2 cross correlation lags at a few selected sites. At Darwin, Port Moresby, Vanimo, Jicamarca, Puerto Rico, Tokyo and Juliusruh there is a significantly reduced lag in Mocha compared to IRI, this improvement is most noticeable at Darwin. At Djibouti, both models perform equally as they both have extremely small lags. In comparison, at Canberra, IRI has almost no lag problem, but Mocha has some issues, for example, in December $S_{12} = 115$ Mocha's foF2 lag is 1.7 hours but IRI's lag is only 0.6 hours. Overall, Mocha has a reduced lag issue in comparison to IRI.

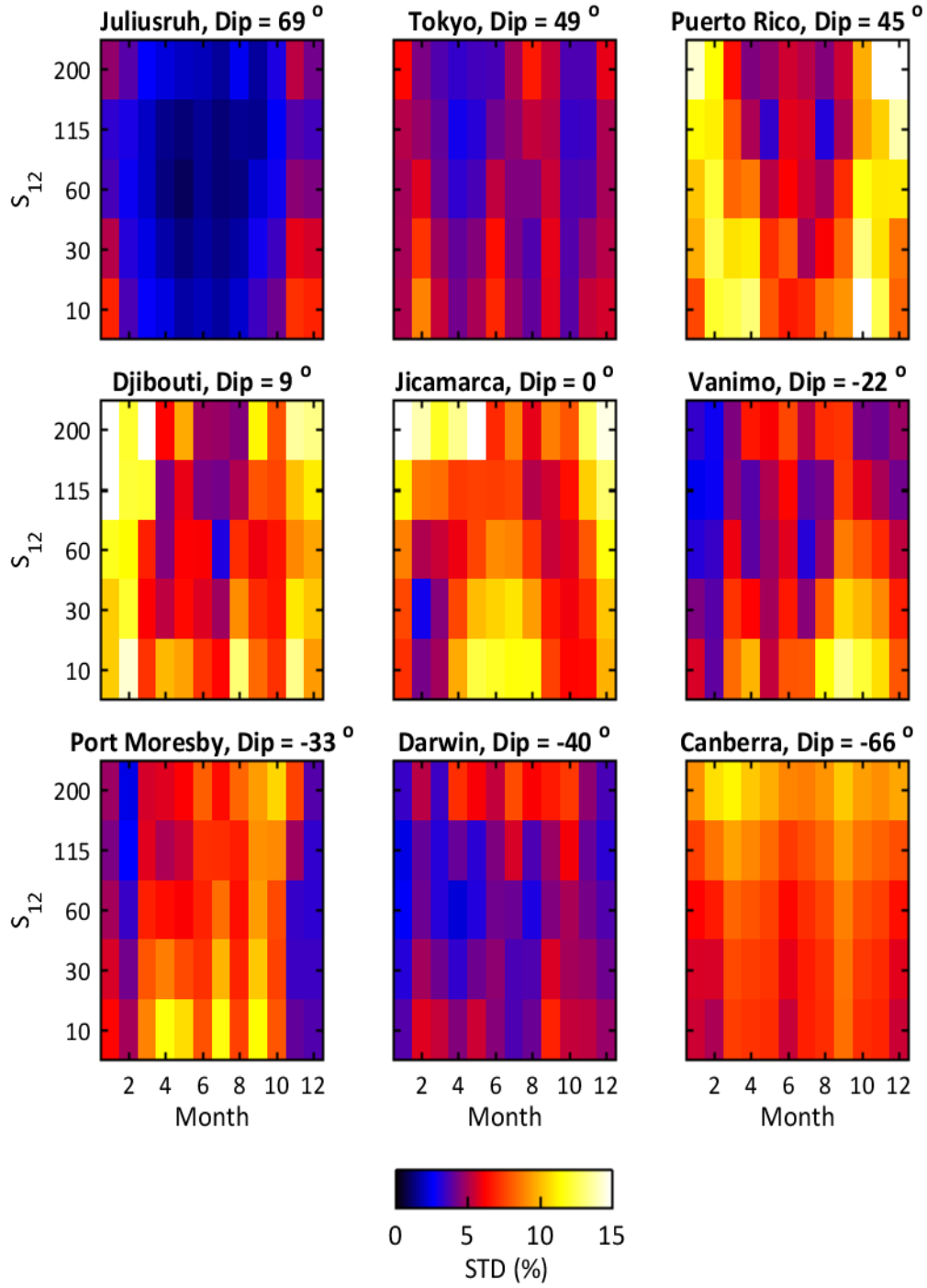

Figure 5.10: The standard deviation of the foF2 residuals between Mocha's fifty percentile foF2 model and monthly median observations for several sites ordered by geomagnetic dip angle. From left to right, Top: Juliusruh (Dip = 69 °), Tokyo (Dip = 49 °), Puerto Rico (Dip = 46 °); middle: Djibouti (Dip = 11 °), Jicamarca (Dip = 1 °), Vanimo (Dip = -22 °); bottom: Port Moresby (Dip = -33 °), Darwin (Dip = -40 °), Canberra (Dip = -66 °).

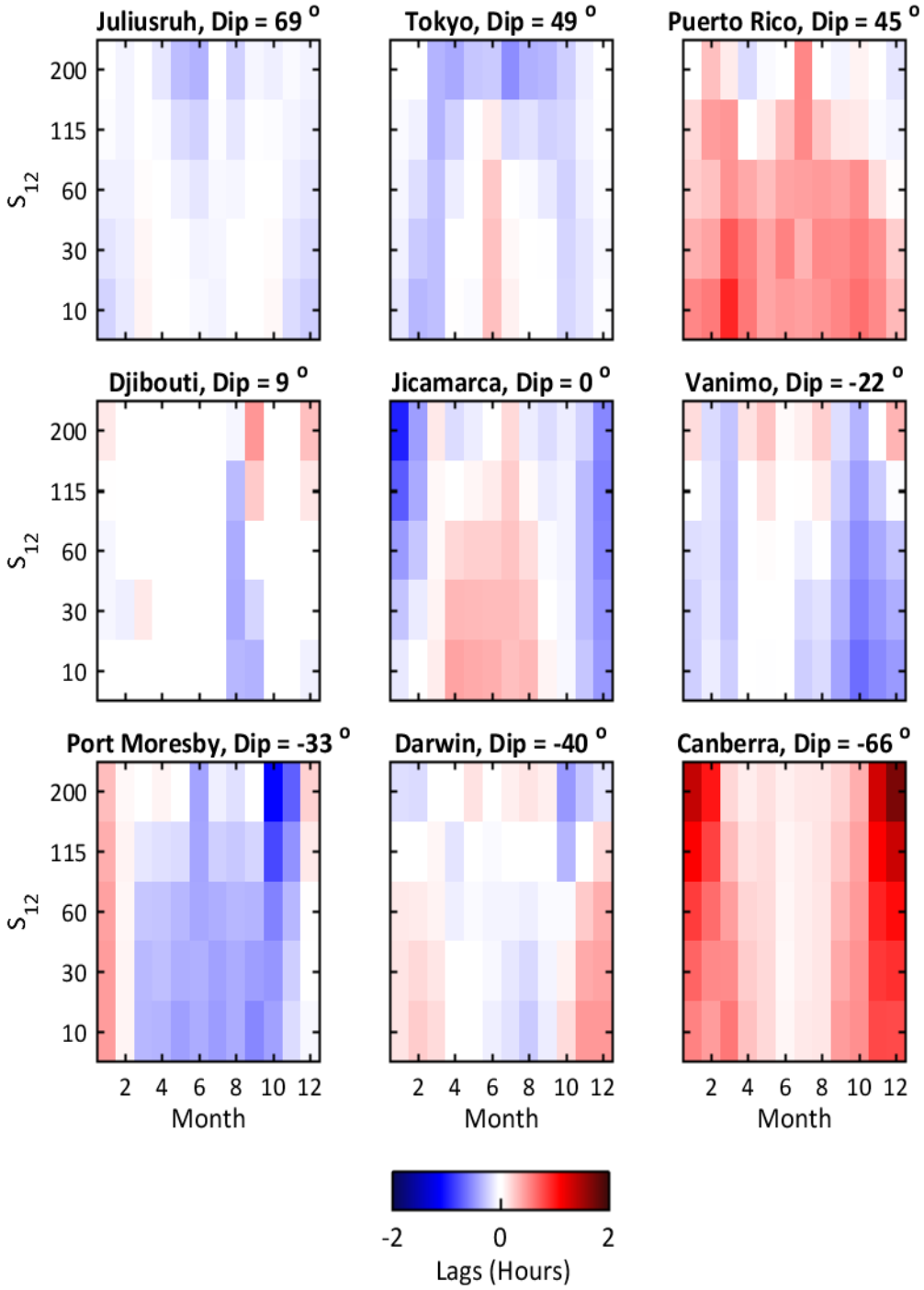

Figure 5.11: The cross correlation lags between Mocha's fifty percentile foF2 model and monthly median observations for several sites ordered by geomagnetic dip angle.. From left to right, Top: Juliusruh (Dip = 69 °), Tokyo (Dip = 49 °), Puerto Rico (Dip = 46 °); middle: Djibouti (Dip = 11 °), Jicamarca (Dip = 1 °), Vanimo (Dip = -22 °); bottom: Port Moresby (Dip = -33 °), Darwin (Dip = -40 °), Canberra (Dip = -66 °).

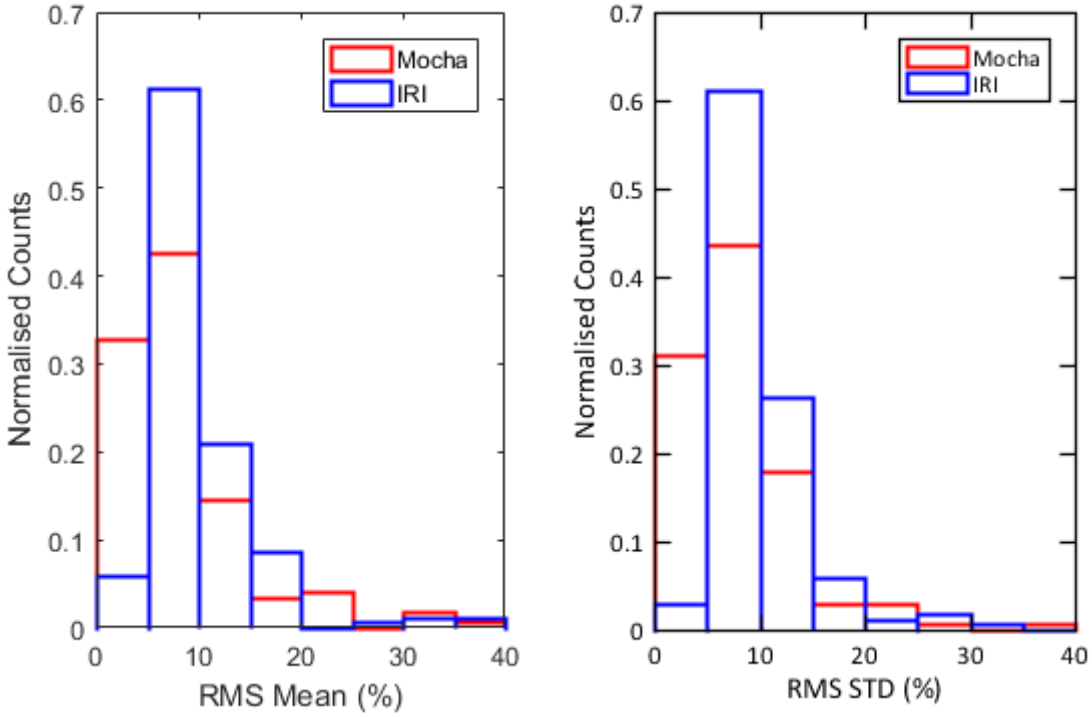

Figure 5.12: The RMS mean (left) and RMS standard deviation (right) of residuals for IRI's median foF2 model (blue line) and Mocha's median foF2 model (red line). Measured using 176 sites.

The distribution of the RMS mean and standard deviation of IRI's and Mocha's median foF2 model for all sites is displayed in Figure 5.12. Overall, it shows that Mocha tends to provide an improvement over IRI in RMS mean and standard deviation. While both Mocha's and IRI's foF2 median models have peaks in the 5 % to 10 % RMS mean bin, Mocha's foF2 median model has many more results in the 0 % to 5 % bin. The RMS standard deviation of residuals also shows that the Mocha's foF2 model outperforms IRI's foF2 model with a greater proportion of standard deviations in the lowest 0 % to 5 % bin.

5.7 foF2 Tenth And Ninetieth Percentile Predictions

Mocha also produces tenth and ninetieth percentile predictions of foF2, the lower and upper deciles respectively. Figures 5.13, 5.14 and 5.15 display the mean and standard deviation of the residuals and lag results (between model and monthly lower decile data) for the ten percentiles predictions, and Figures 5.16, 5.17 and 5.18 displays the same results for the ninetieth percentile model. The mean and standard deviation of the residuals and lag results for the ten, fifty and ninetieth

percentile are similar.

Figure 5.19 displays histograms of the RMS mean and standard deviation of the residuals for the tenth and ninetieth percentiles together with the median measured at 140 sites. It is difficult to quantify which model performs better; they appear to perform equally well as a majority of the sites are within the first three bins, between 0 % to 15 %.

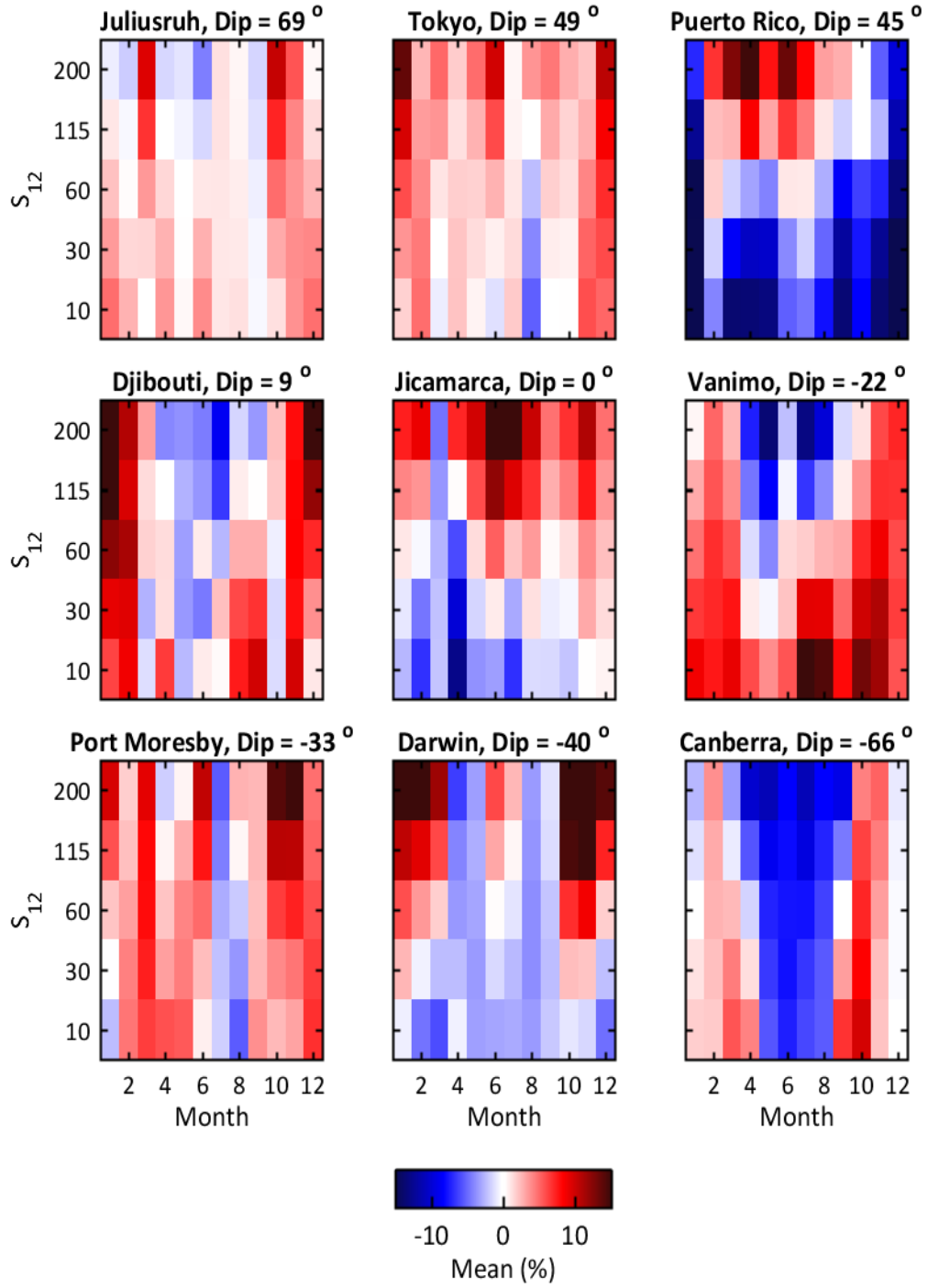

Figure 5.13: The mean foF2 residual between Mocha's ten percentile foF2 model and monthly ten percentile observations. From left to right, Top: Juliusruh (Dip = 69 °), Tokyo (Dip = 49 °), Puerto Rico (Dip = 46 °); middle: Djibouti (Dip = 11 °), Jicamarca (Dip = 1 °), Vanimo (Dip = -22 °); bottom: Port Moresby (Dip = -33 °), Darwin (Dip = -40 °), Canberra (Dip = -66 °) sorted by geomagnetic dip angle.

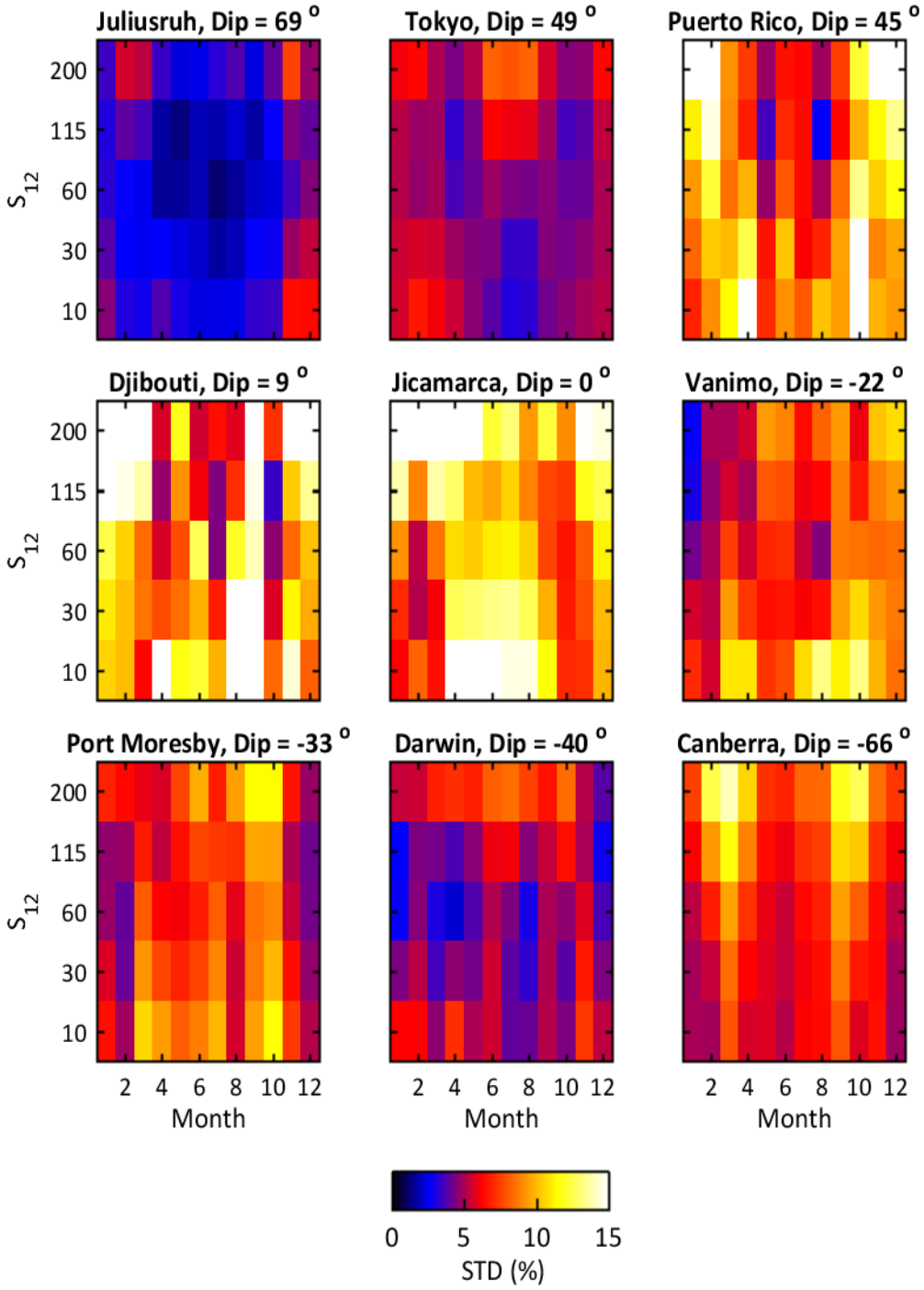

Figure 5.14: The standard deviation of the foF2 residuals between Mocha's ten percentile foF2 model and ten percentile observations. From left to right, Top: Juliusruh (Dip = 69 °), Tokyo (Dip = 49 °), Puerto Rico (Dip = 46 °); middle: Djibouti (Dip = 11 °), Jicamarca (Dip = 1 °), Vanimo (Dip = -22 °); bottom: Port Moresby (Dip = -33 °), Darwin (Dip = -40 °), Canberra (Dip = -66 °) sorted by geomagnetic dip angle.

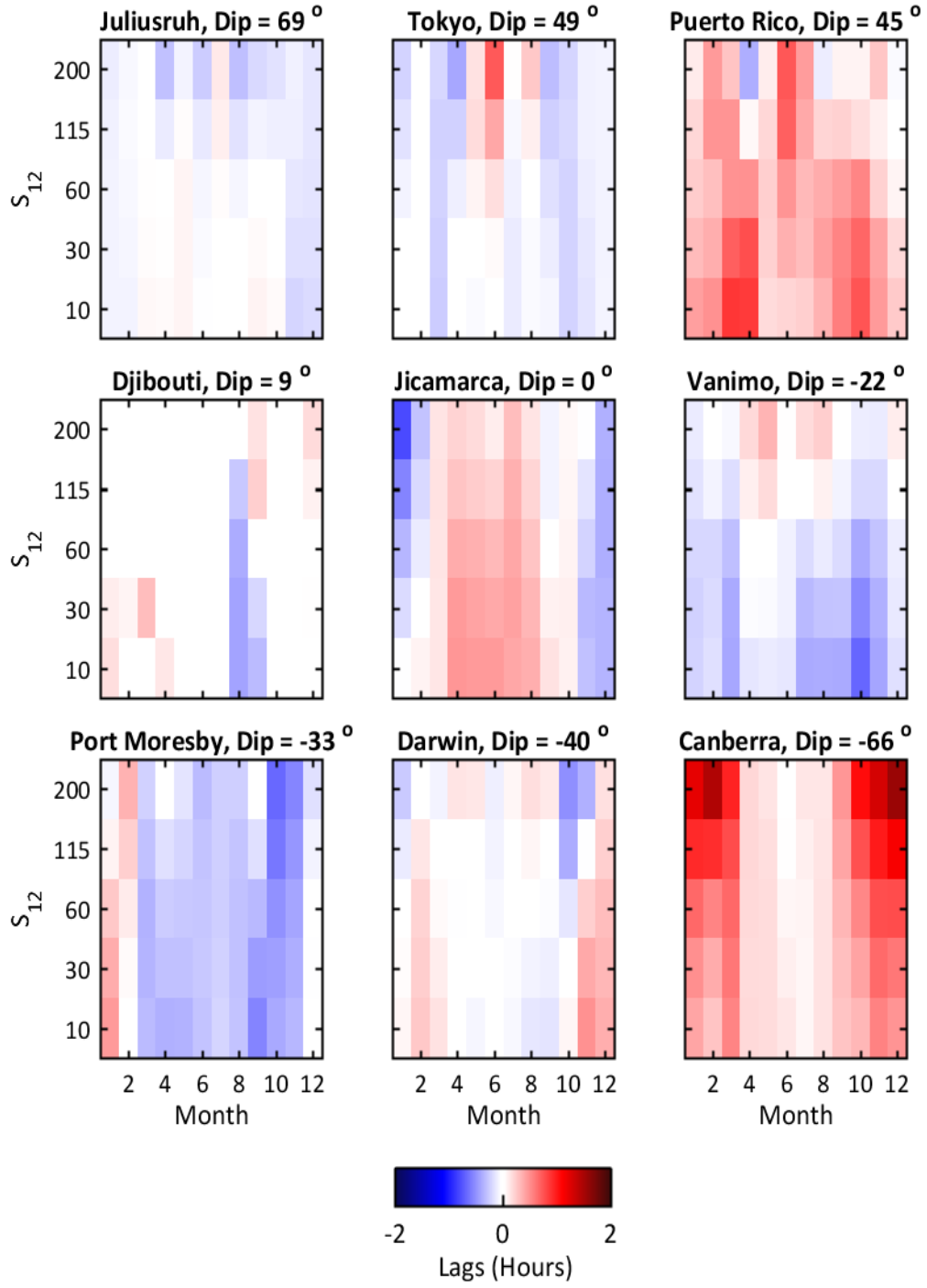

Figure 5.15: The cross correlation lags between Mocha's ten percentile foF2 model and ten percentile observations. From left to right, Top: Juliusruh (Dip = 69 °), Tokyo (Dip = 49 °), Puerto Rico (Dip = 46 °); middle: Djibouti (Dip = 11 °), Jicamarca (Dip = 1 °), Vanimo (Dip = -22 °); bottom: Port Moresby (Dip = -33 °), Darwin (Dip = -40 °), Canberra (Dip = -66 °) sorted by geomagnetic dip angle.

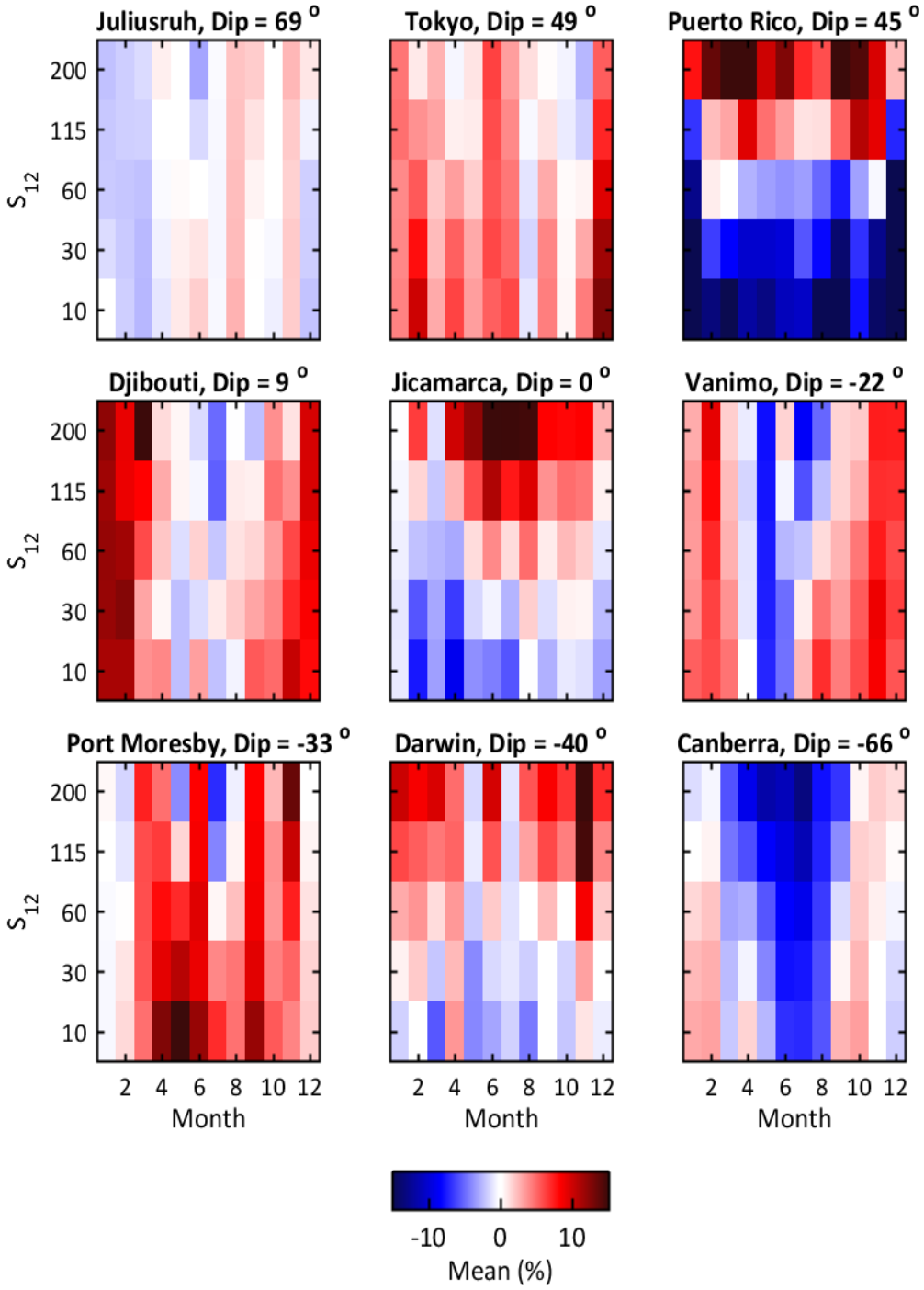

Figure 5.16: The mean foF2 residual between Mocha's ninety percentile foF2 model and ninety percentile observations. From left to right, Top: Juliusruh (Dip = 69 °), Tokyo (Dip = 49 °), Puerto Rico (Dip = 46 °); middle: Djibouti (Dip = 11 °), Jicamarca (Dip = 1 °), Vanimo (Dip = -22 °); bottom: Port Moresby (Dip = -33 °), Darwin (Dip = -40 °), Canberra (Dip = -66 °) sorted by geomagnetic dip angle.

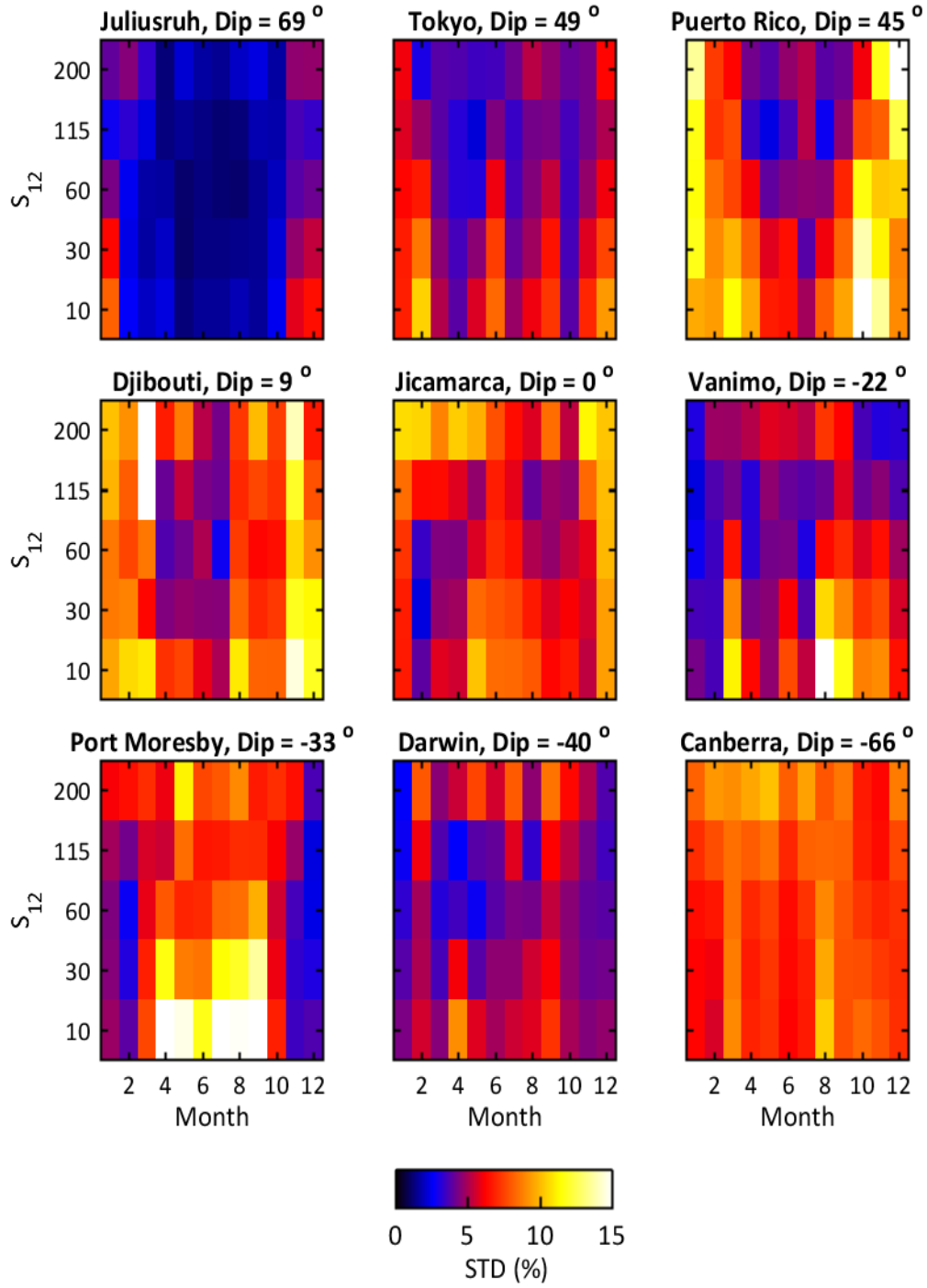

Figure 5.17: The standard deviation of the foF2 residuals between Mocha's ninety percentile foF2 model and ninety percentile observations. From left to right, Top: Juliusruh (Dip = 69 °), Tokyo (Dip = 49 °), Puerto Rico (Dip = 46 °); middle: Djibouti (Dip = 11 °) , Jicamarca (Dip = 1 °), Vanimo (Dip = -22 °); bottom: Port Moresby (Dip = -33 °), Darwin (Dip = -40 °), Canberra (Dip = -66 °) sorted by geomagnetic dip angle.

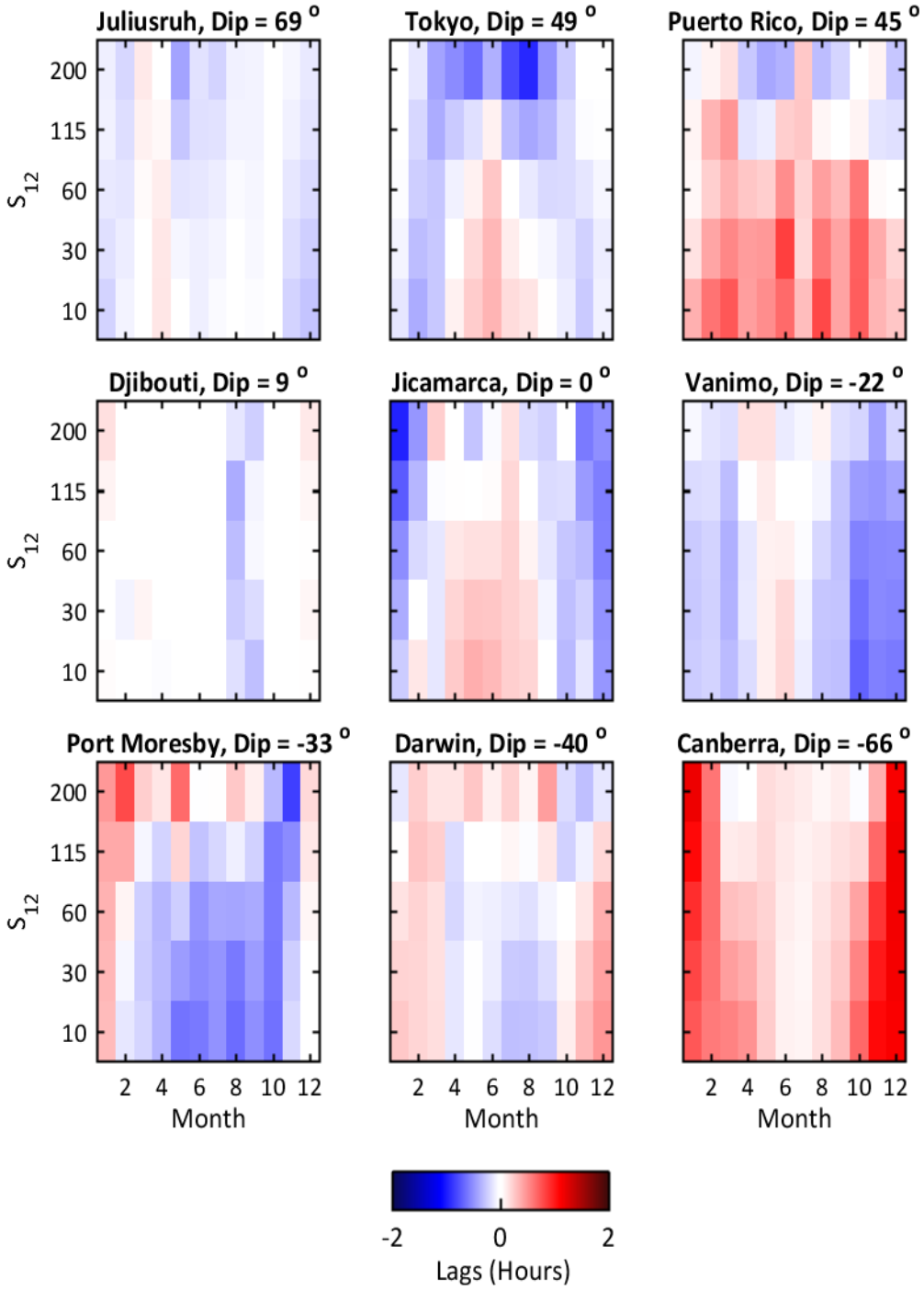

Figure 5.18: The cross correlation lags between Mocha's ninety percentile foF2 model and ninety percentile observations. From left to right, Top: Juliusruh (Dip = 69 °), Tokyo (Dip = 49 °), Puerto Rico (Dip = 46 °); middle: Djibouti (Dip = 11 °), Jicamarca (Dip = 1 °), Vanimo (Dip = -22 °); bottom: Port Moresby (Dip = -33 °), Darwin (Dip = -40 °), Canberra (Dip = -66 °) sorted by geomagnetic dip angle.

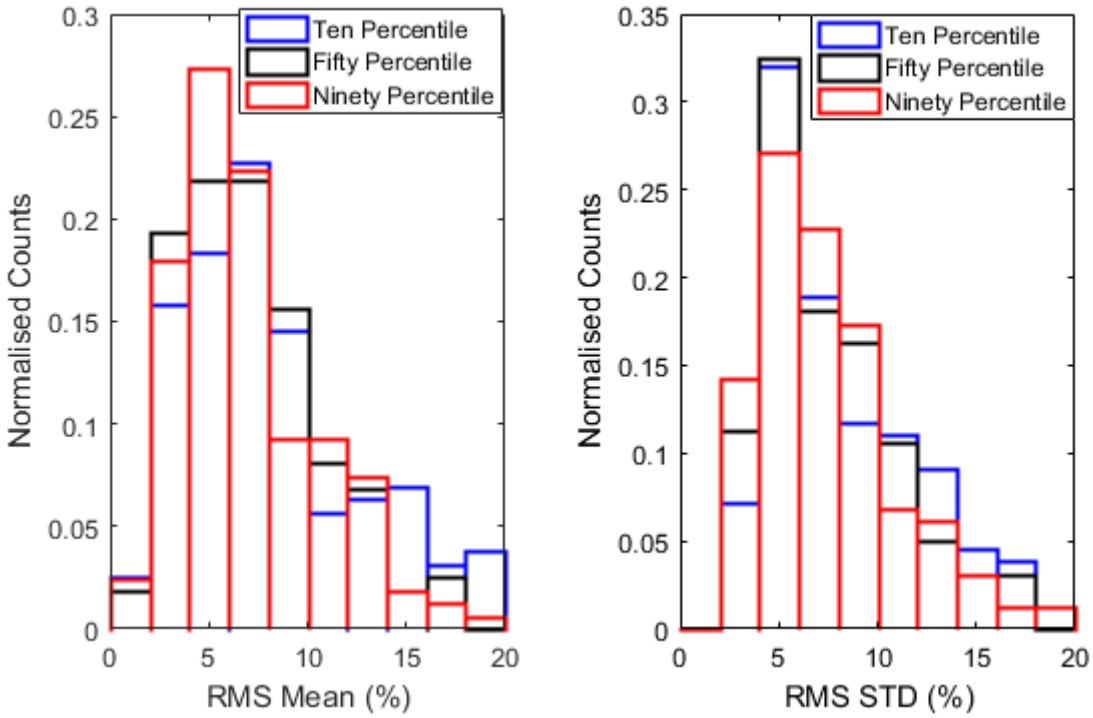

Figure 5.19: The RMS mean (left) and RMS standard deviation (right) of the residuals for Mocha's ten percentile foF2 model (blue line), Mocha's fifty percentile foF2 model (black line) and Mocha's ninety percentile foF2 model (red line). Measured using 176 sites.

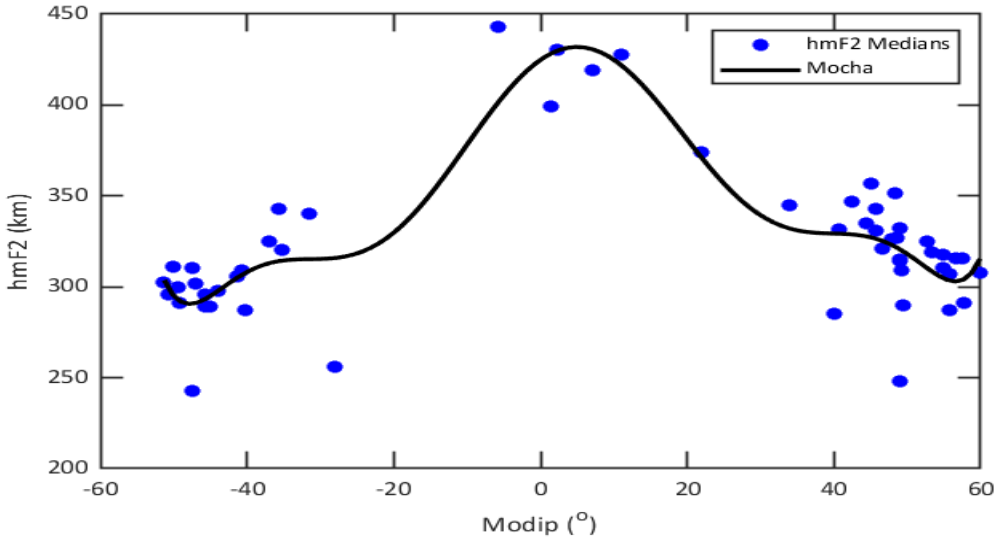

Figure 5.20: Mocha's hmF2 fit (black line) at 14 LT, May, $S_{12} = 200$. There are 54 observations and Mocha uses an 8th order Chebychev fit.

5.8 Mocha's hmF2 model

Mocha also produces global empirical monthly median and variance hmF2 predictions using the same method as the foF2 model, but the maximum order fit used is 8 as there are fewer data available. Figure 5.20 displays an example of Mocha's fits for 14 LT, in May at high solar activity ($S_{12} = 200$). This month, year and solar cycle combination was chosen to display the complicated low latitude features in the observations. A majority of the data comes from the mid latitude regions. There is also some data near the magnetic equator at $\sim 0\,^\circ$ modified dip angle. Unfortunately there are no data in the Southern Hemisphere between -20 ° to -10 ° modified dip angle and there is only one site in the Northern Hemisphere in the 10 ° to 20 ° modified dip angle range. The limited data availability makes curve fitting in these sub-equatorial regions difficult, and considerable extrapolation is required.

It is possible to compare IRI's and Mocha's median hmF2 model at Curtin (Lat = -18 °N, Long = 124 °E, Dip = -48 °), noting that IRI does not use Curtin data within the model but Mocha does. Curtin was chosen as it was one of the few Australian sites with a solar cycle of hmF2 observations. Figures 5.21, 5.22 and 5.23 display IRI's and Mocha's median hmF2 mean, standard deviation residuals and lag respectively. It is clear that Mocha tends to perform fairly well, with large improvements over IRI in the standard deviation of residuals and lag.

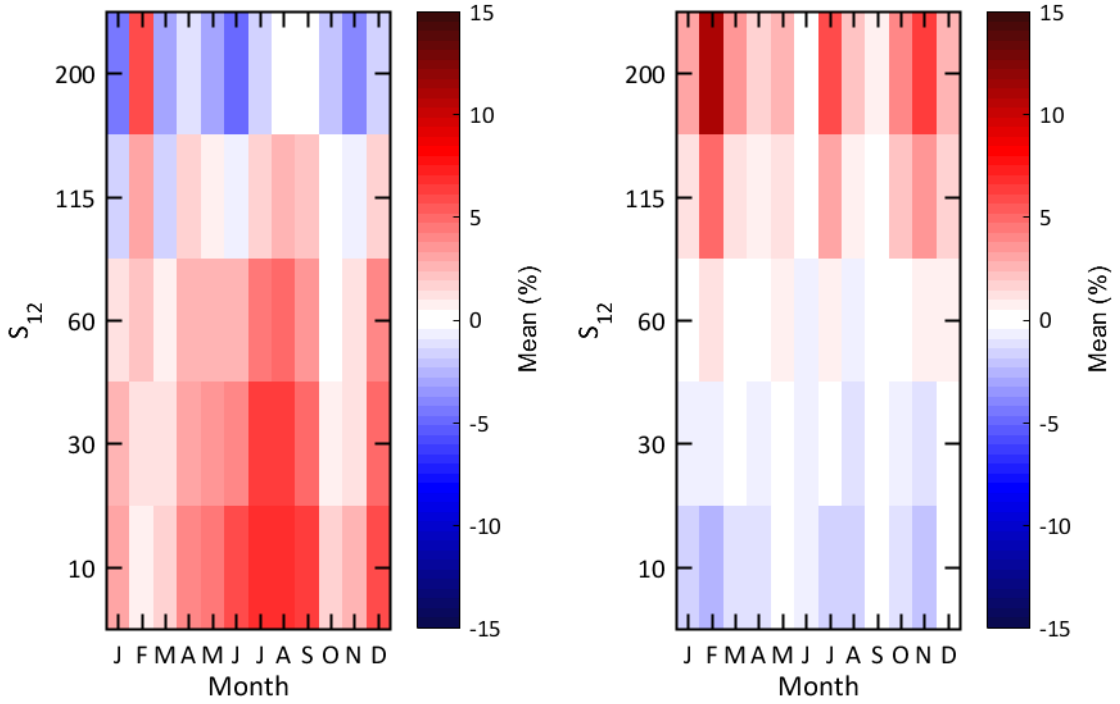

Figure 5.21: A comparison of the mean residual in the monthly median hmF2 model at Curtin, with IRI (left) and Mocha (right).

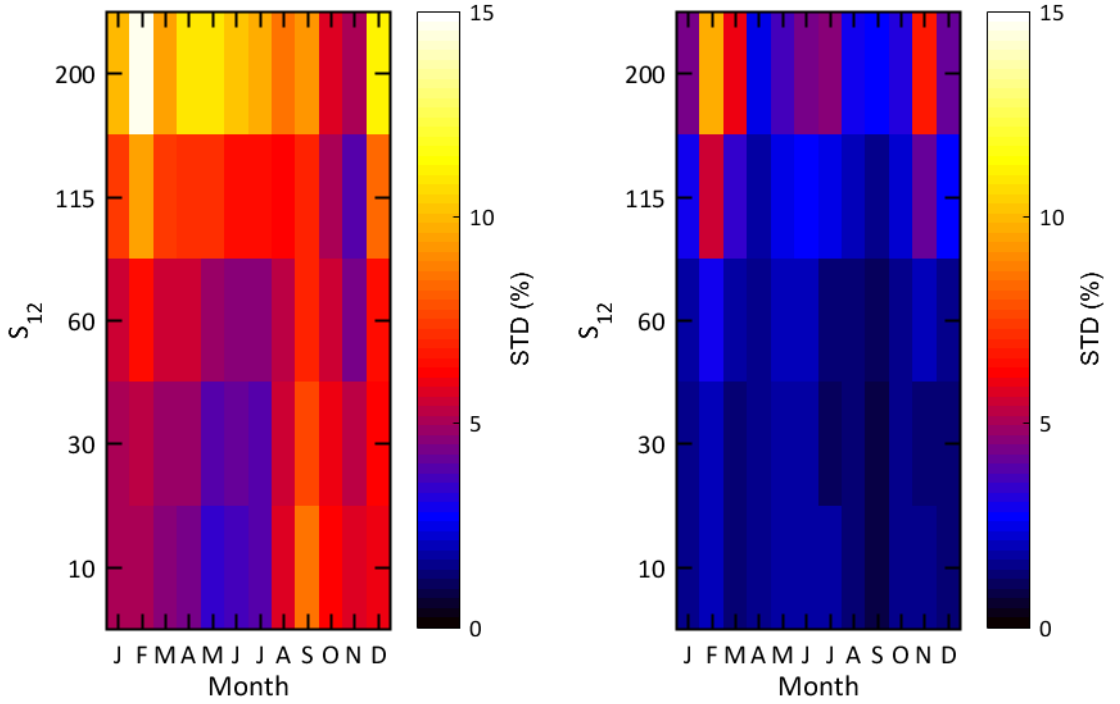

Figure 5.22: A comparison of the standard deviation of the residuals in the monthly median hmF2 model at Curtin, with IRI (left) and Mocha (right).

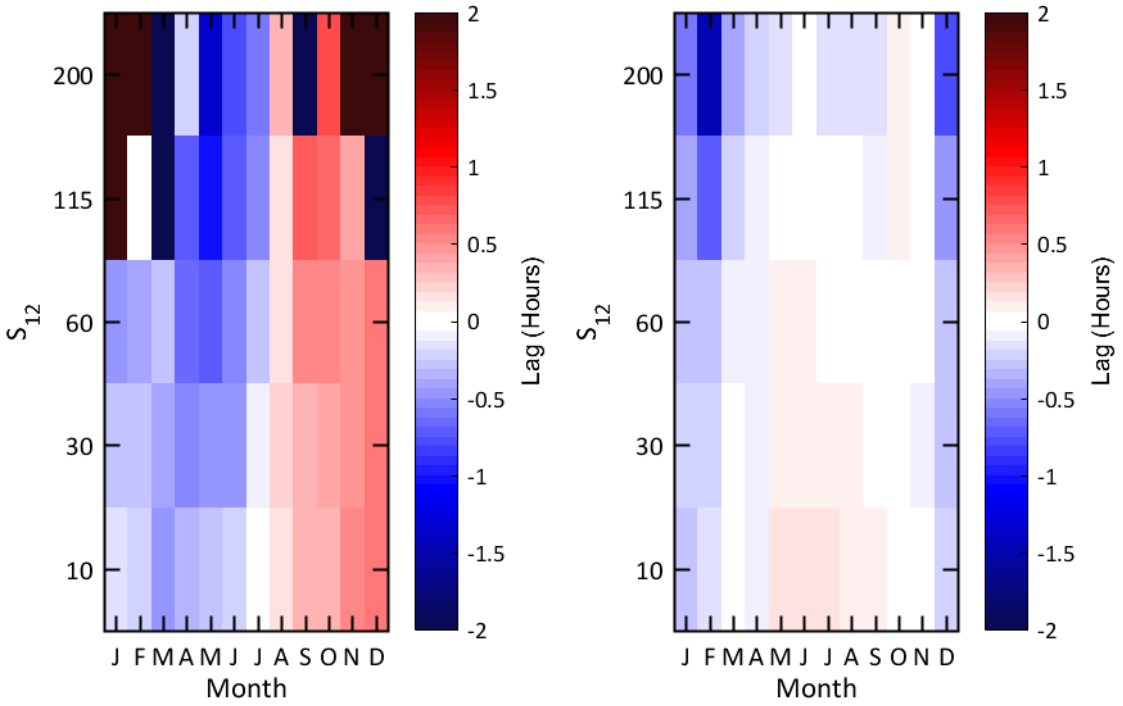

Figure 5.23: A comparison of the cross correlation lag in the monthly median hmF2 model at Curtin, with IRI (left) and Mocha (right).

This was repeated at Juliusruh with the mean, standard deviation and lag results in Figures 5.24, 5.25 and 5.26. Again there is a clear improvement in Mocha over IRI in every month at Juliusruh, particularly in the mean and standard deviations.

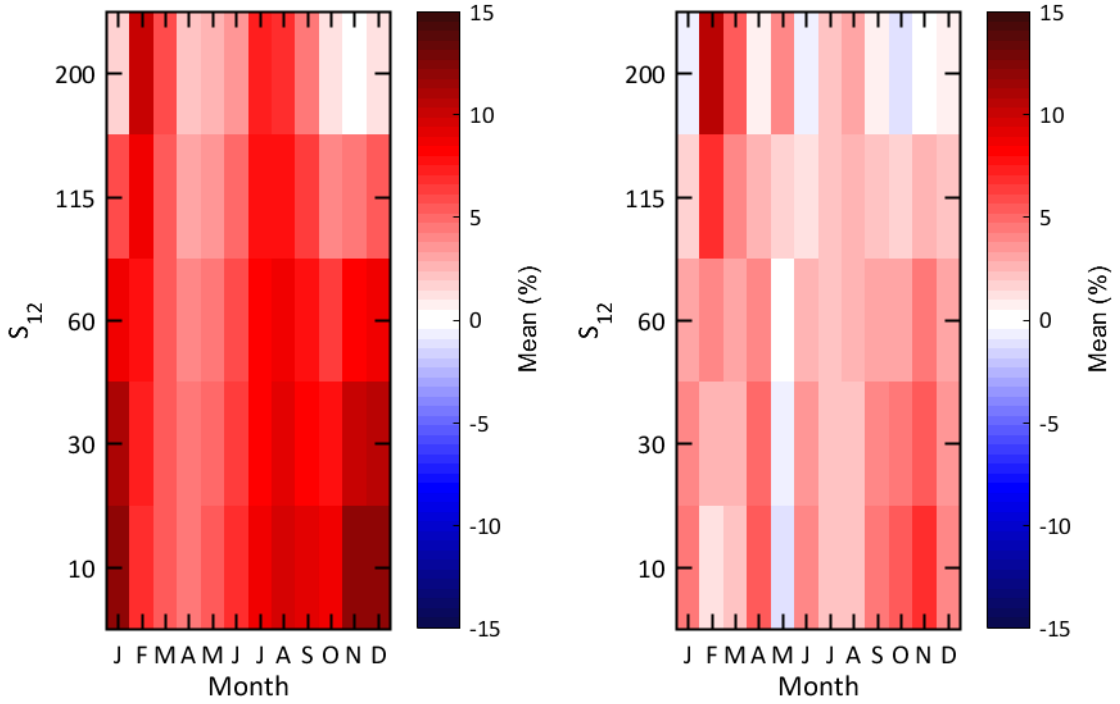

Figure 5.24: A comparison of the mean residual in the monthly median hmF2 model at Juliusruh, with IRI (left) and Mocha (right).

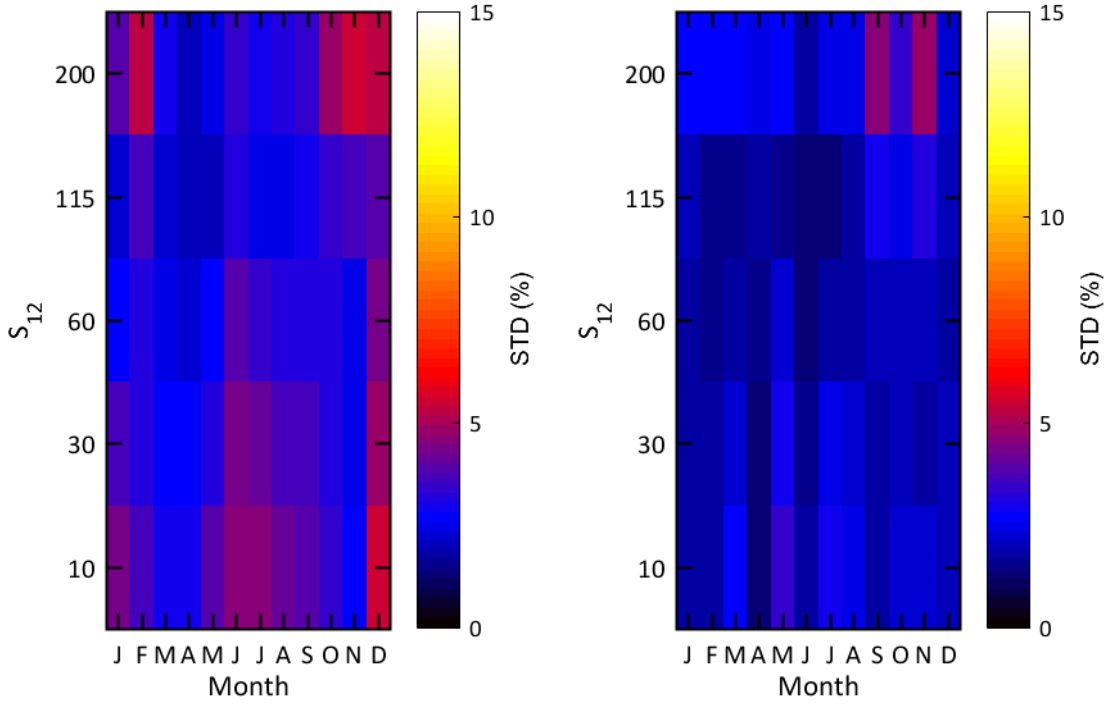

Figure 5.25: A comparison of the standard deviation of the residuals in the monthly median hmF2 model at Juliusruh, with IRI (left) and Mocha (right).

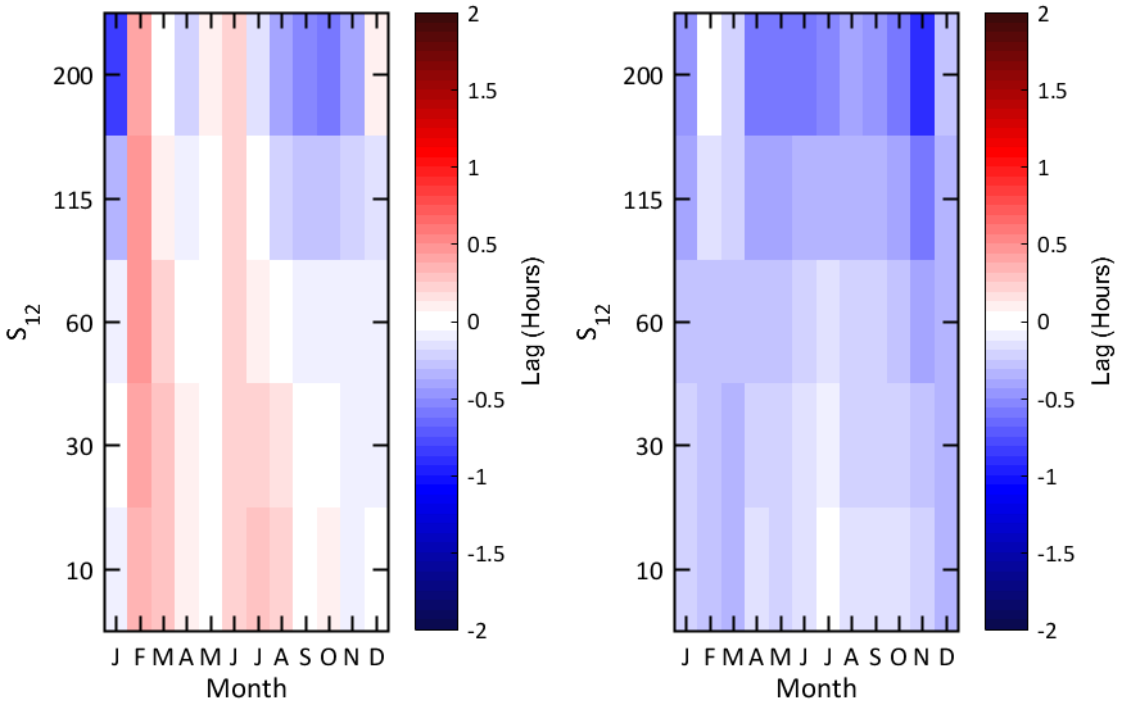

Figure 5.26: A comparison of the cross correlation lag in the monthly median hmF2 model at Juliusruh, with IRI (left) and Mocha (right).

IRI's and Mocha's hmF2 monthly median model was tested at 75 sites, with the RMS mean and RMS standard deviation of residuals calculated at each site. Histograms of these data are displayed in Figure 5.27. There is a large improvement in Mocha's hmF2 model over IRI's model as Mocha peaks in lower bias and standard deviation bins than IRI.

Mocha's hmF2 tenth and ninetieth percentile model RMS mean and standard deviation of residuals were also calculated and are displayed in Figure 5.28 along with the earlier fiftieth percentile results. It is clear that the tenth and ninetieth percentile models perform equally well as the median model.

5.9 Global hmF2 Analysis

A global review of hmF2 was performed by sorting a few selected sites by geomagnetic dip angle and producing the solar cycle summary plots at each site. The mean, standard deviation of the residuals are displayed in Figures 5.29 and 5.30, with lag measurements displayed in Figure 5.31.

Examination of the mean residual in Mocha's hmF2 model in Figure 5.29 shows the mean error is extremely low with the exception of Darwin, solar maximum,

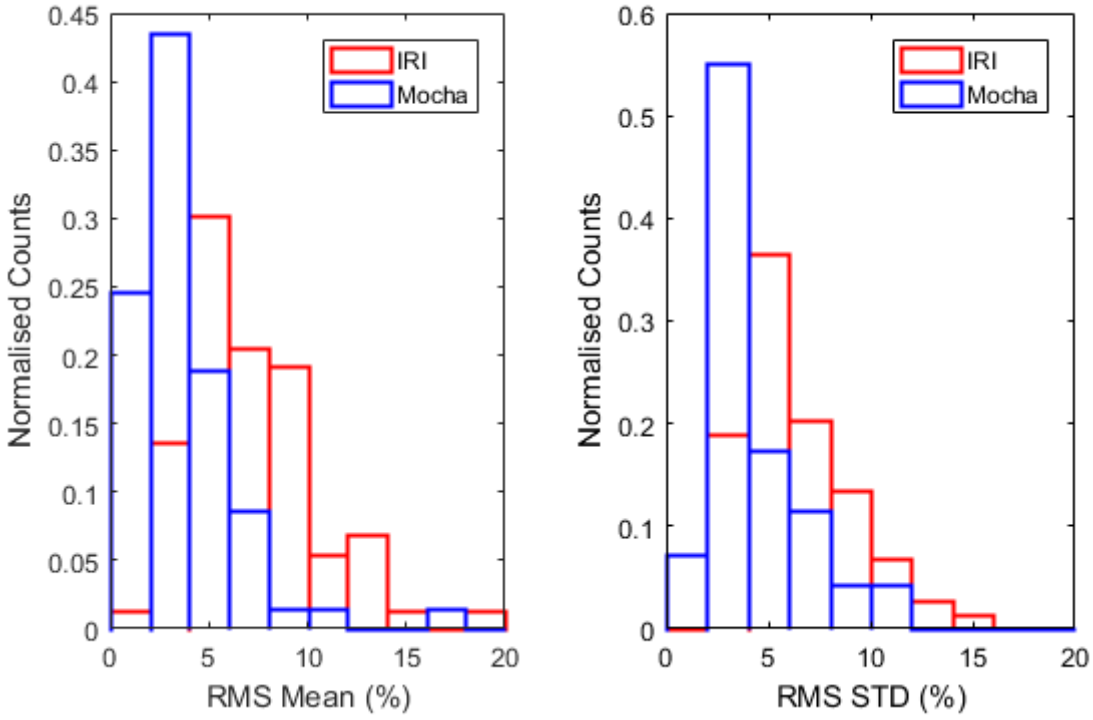

Figure 5.27: The RMS mean (left) and RMS standard deviation (right) of residuals for IRI's monthly median hmF2 (red line) and Mocha's monthly median hmF2 (blue line) measured at 75 sites.

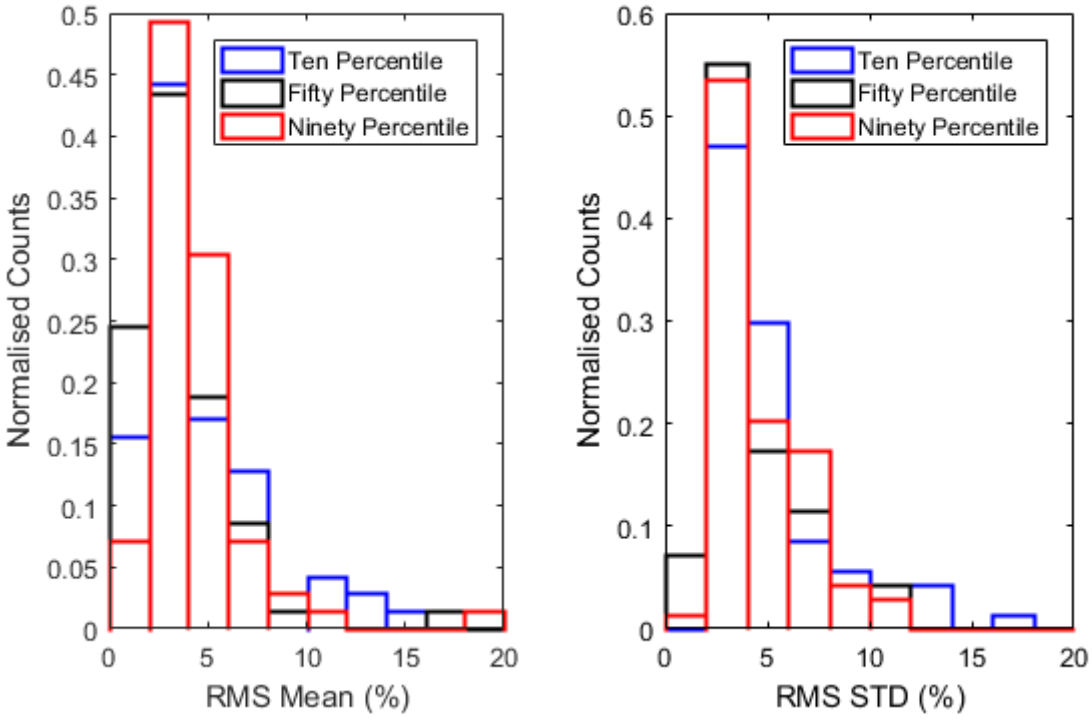

Figure 5.28: The RMS mean (left) and RMS standard deviation (right) for Mocha's ten percentile (blue line), median (black line) and ninety percentile (red line) model residuals calculated at 75 sites.

where Mocha considerably underestimates hmF2. In general the mean residual is solar cycle independent with the exception of Darwin and Jeju, and to a slight extent Puerto Rico. Season appears to be the dominating factor at Sanya, Jicamarca and Ajana. At Ajana and Sanya, Mocha underestimates in winter. At Jicamarca, Mocha underestimates in November to February and overestimates in May. Mocha's mean residual appears to do best in the Southern Hemisphere.

Mocha's median hmF2 model standard deviation of residuals results are displayed in Figure 5.30. Mocha does extremely well overall with the exception of Darwin. Ajana, Juliusruh, Jeju, Scherger, and Sanya have a relatively constant standard deviation at $\sim 1\,\%$ to $4\,\%$. At Puerto Rico, Guam and Jicamarca the standard deviation typically reaches up to $\sim 10\,\%$. Darwin is the exception site with a few saturated ($> 15\,\%$) standard deviations at high sunspot numbers. The standard deviations are fairly seasonal and solar cycle independent, with the exception of Puerto Rico and Darwin which have larger standard deviations in solar maximum. It appears that Mocha's standard deviations are worst during the mid to low latitude boundary.

The cross correlation lags in Mocha's hmF2 median model displayed in Figure 5.31 are fairly minimal. There appears to be some positive lags at Jicamarca in summer and some negative lags in Puerto Rico at high sunspot numbers. Overall, there does not appear to be a serious lag problem within the model.

Mocha's hmF2 median model mean residual displayed in Figure 5.29 may be compared with IRI's hmF2 median mean residuals in Figure 4.43. It is clear that IRI tends to significantly overestimate hmF2 at almost every site except Ajana. By comparison Mocha, has an improved mean at every site except Darwin, solar maximum, where IRI and Mocha perform relatively equally.

Comparison of IRI's and Mocha's hmF2 median model standard deviations of residuals in Figures 4.44 and 5.30, shows that again Mocha provides a large improvement over IRI at each site except at Jicamarca where IRI performs slightly better. The most significant improvements are at the low latitudes, with minimal improvements by Mocha over IRI at the mid latitudes.

Finally, the cross correlation lags of Mocha's and IRI's hmF2 models can be compared from examination of Figures 5.31 and 4.48. Mocha provides a clear improvement to the lag results compared to IRI. This improvement is particularly striking at the low latitude sites Sanya, Guam, Scherger and Darwin where the cross correlation lag results have extreme positive or negative values for IRI but with small lag values in Mocha.

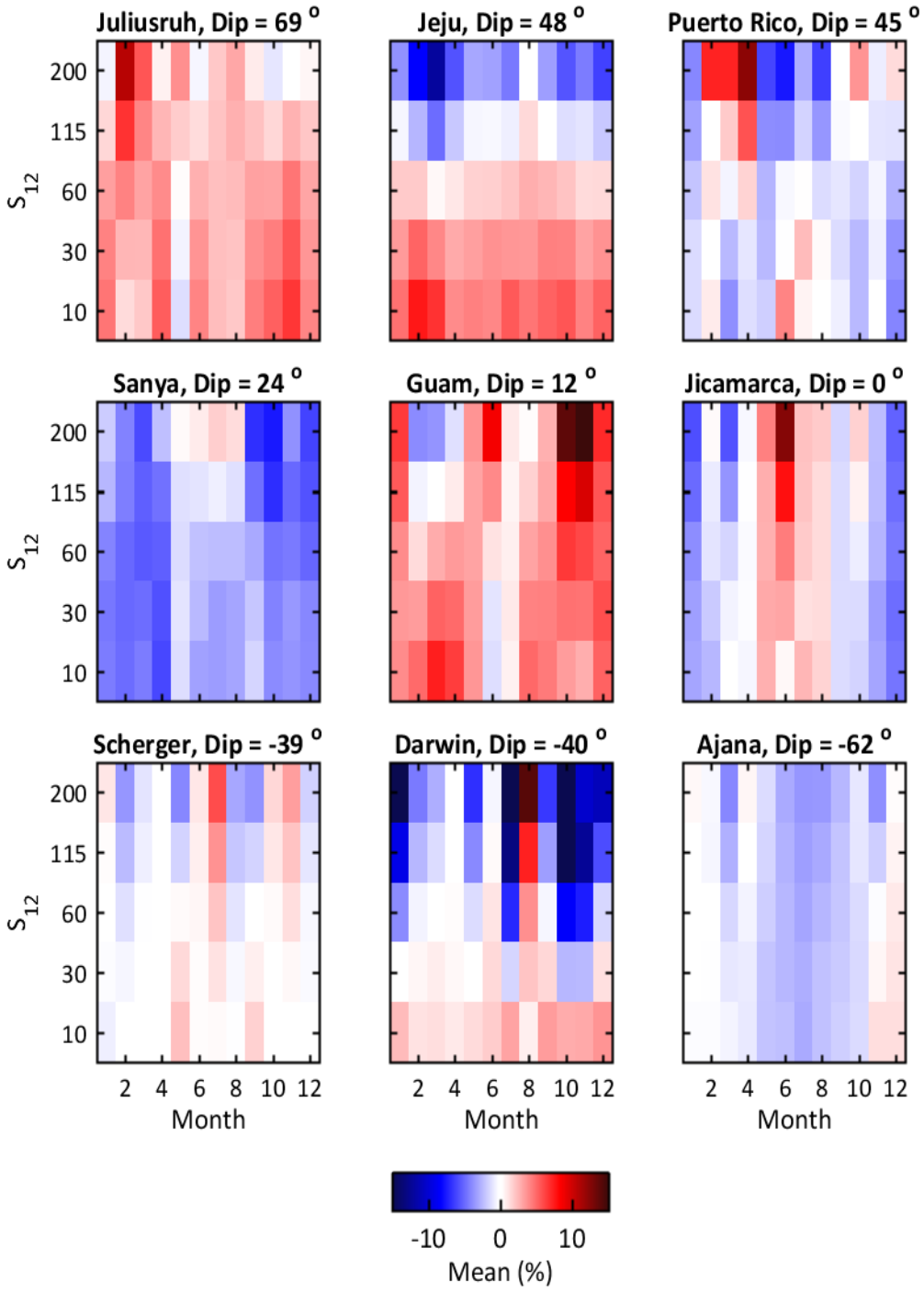

Figure 5.29: The mean hmF2 residual from Mocha's hmF2 median model. From left to right, Top: Juliusruh (Dip = 69 °), Jeju (Dip = 48 °), Puerto Rico (Dip = 46 °), Middle: Sanya (Dip = 24 °), Guam (Dip = 12 °), Jicamarca (Dip = 1 °); Bottom: Scherger (Dip = -40 °), Darwin (Dip = -40 °), Ajana (Dip = -63 °) ordered by geomagnetic dip angle.

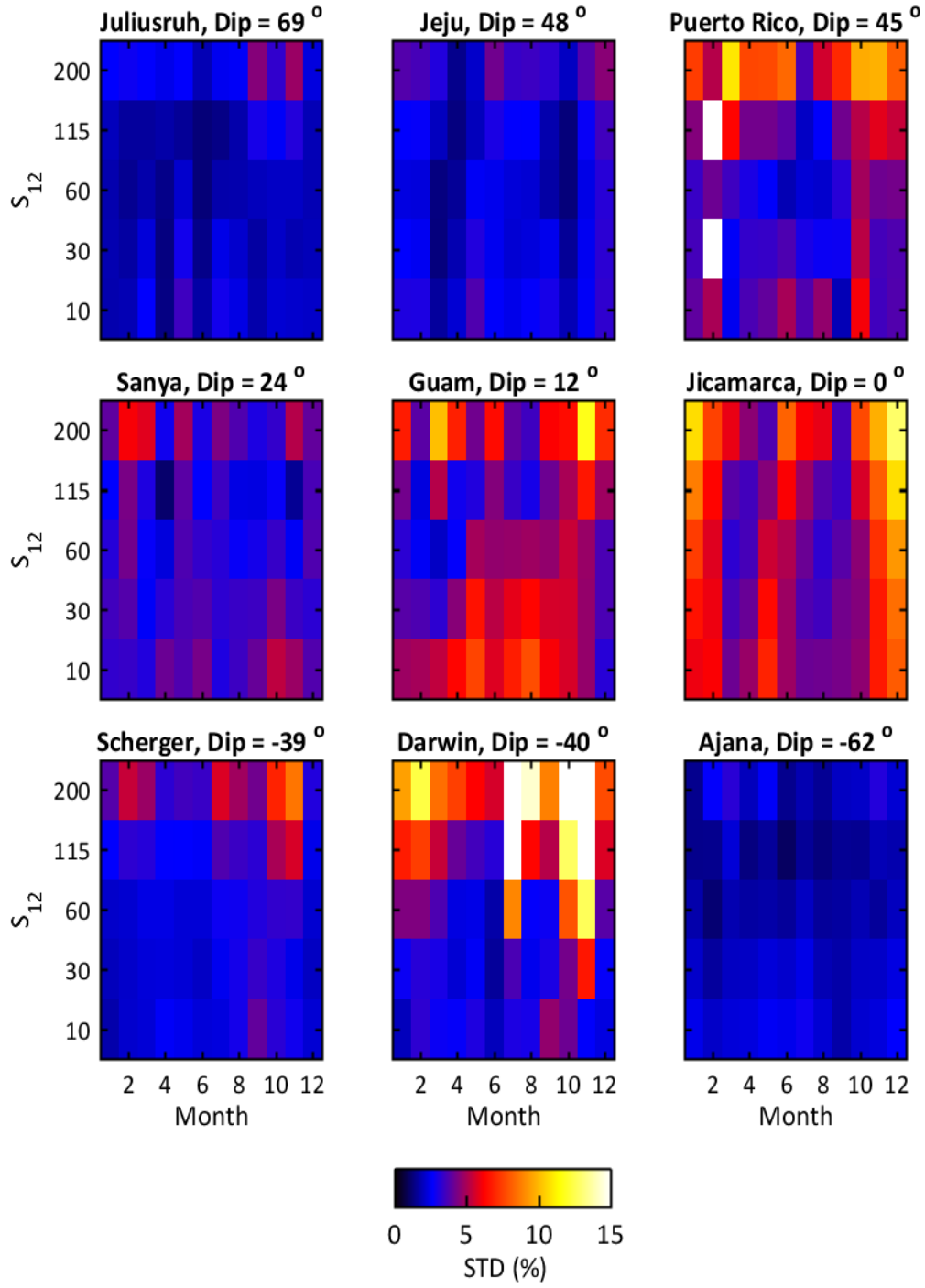

Figure 5.30: The standard deviation hmF2 residual from Mocha's hmF2 median model. From left to right, Top: Juliusruh (Dip = 69 °), Jeju (Dip = 48 °), Puerto Rico (Dip = 46 °), Middle: Sanya (Dip = 24 °), Guam (Dip = 12 °), Jicamarca (Dip = 1 °); Bottom: Scherger (Dip = -40 °), Darwin (Dip = -40 °), Ajana (Dip = -63 °) ordered by geomagnetic dip angle.

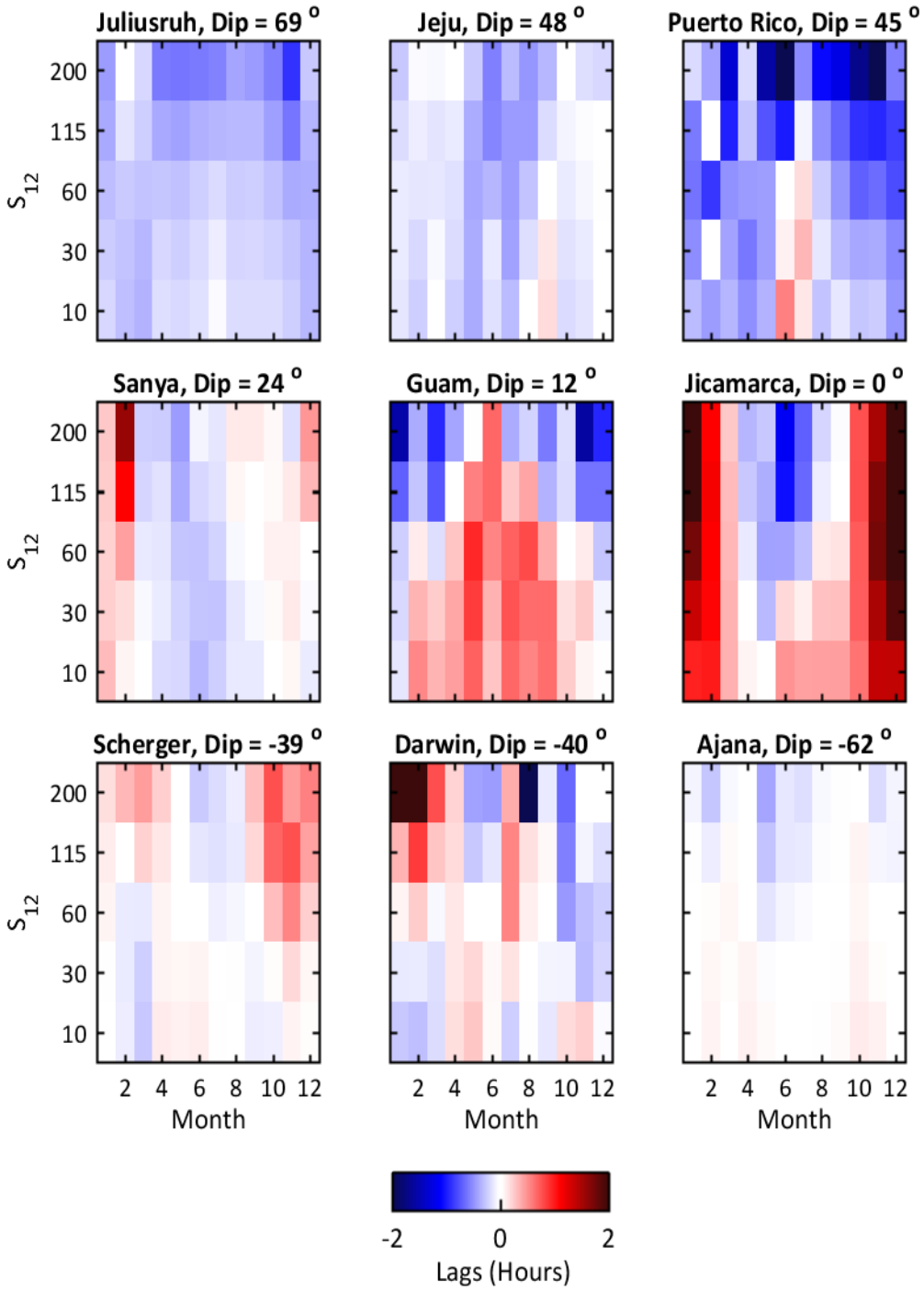

Figure 5.31: The cross correlation lags in hmF2 from Mocha's hmF2 median model. From left to right, Top: Juliusruh (Dip = 69 °), Jeju (Dip = 48 °), Puerto Rico (Dip = 46 °), Middle: Sanya (Dip = 24 °), Guam (Dip = 12 °), Jicamarca (Dip = 1 °); Bottom: Scherger (Dip = -40 °), Darwin (Dip = -40 °), Ajana (Dip = -63 °) ordered by geomagnetic dip angle.

5.10 Summary

Three methodologies were investigated to produce an accurate model of foF2. The initial version, Decaf, averaged the medians from the surrounding sites. This technique was severely limited due to a lack of data and an inability to provide a latitude correction.

Latte, provides a significant improvement over Decaf by using polynomial fits on median observations across latitude, performed for each hour for each month for each year. However this had difficulties as the quality of fits would vary dramatically in each month and hour.

Mocha develops Latte's methodology further to make the fitting more robust by binning data into five sunspot regimes, $S_{12} = [0, 20]$, $[20, 40]$, $[40, 80]$, $[80, 150]$ and $[150, 250]$, and then using linear interpolation to generate fits at $S_{12} = 10$, 30, 60, 115, and 200.

The comparison between IRI's hmF2 median model and Mocha's hmF2 median model shows that Mocha provides a large improvement. While IRI tends to overestimate hmF2, Mocha's hmF2 mean residual is within a few percent. Similarly, the standard deviations indicate that Mocha provides an improvement at low latitudes. The RMS results support this as the Mocha results tends to be in the lower histogram bins compared to IRI.

Mocha's hmF2 model is robust with its predictions performing equally well at the tenth, median and ninetieth percentiles. Similar results can be seen in the foF2 tenth, fiftieth and ninetieth percentile models, where each perform equally well.

Overall, Mocha provides a large improvement over IRI's hmF2 predictions and a small improvement over IRI's foF2 predictions. The variance predictions of these two parameters are also robust.

Chapter 6

Conclusion

IRI is a global empirical monthly median model of ionospheric parameters. In the late 1960s, Karl Rawer chaired the IRI project as a joint project between COSPAR and URSI. IRI has been constantly updated since with new techniques and observations via a COSPAR working group. IRI's foF2 model uses a set of global fits at sunspot numbers $R_{12} = 0$ and 100, to generate foF2 predictions, while IRI's hmF2 model methodology is similar.

Mocha is a new global empirical climatological monthly median and variance model of foF2 and hmF2. It uses three Chebychev fits to the median, tenth and ninetieth percentile observation of each parameter as a function of modified dip at each month and at $S_{12} = 10$, 30, 60, 115 and 200 for 176 sites for foF2 and 75 sites for hmF2. It uses linear interpolation between these fits to predict foF2 and hmF2 values for any sunspot number.

As Mocha is a climatological model, short period anomalies need to be removed. The observations were geomagnetic storm filtered by removing days based off Dst thresholds. If the storm was fairly weak (Dst $\leq$-75 nT) the strength of the ionosphere could be expected to recover in a day, so only this day of observations were removed. If the the storm was strong (Dst $\leq$-100 nT) then, as the ionosphere would take longer to recover, the subsequent day of observations was removed as well.

To streamline the analysis, the data were grouped into 5 sunspot bins, $S_{12} =$ [0, 20], [20, 40], [40, 80], [80 150] and [150, 250]. The median observation in each bin was then linearly corrected to $S_{12} = 10$, 30, 60, 115, and 200.

Mocha performs a Chebychev fit on the interpreted S_{12} observations across modified geomagnetic dip angle to predict foF2 and hmF2. It dynamically selects the order of the fit based on minimising the RMSE error of the fits, selecting up to a 14th order polynomial for foF2 and 8th order for hmF2.

The analysis investigated the diurnal, seasonal and solar cycle influence on the residuals in IRI's foF2 model. IRI typically underestimates foF2 during solar minimum and overestimates during solar maximum. In general the foF2 model performs

better in mid latitudes than in low latitudes. IRI's foF2 model also contains a temporal lag, that tends to be positive (up to $\sim$1 hour) in summer (Northern and Southern Hemisphere) with IRI leading the data and negative ($\sim$-0.5 to 0 hours) in winter (Northern and Southern Hemisphere), primarily affecting Southern Hemisphere sites such as Canberra and Darwin. This lag was corrected using a simple temporal adjustment which provided a significant improvement to the standard deviation of the foF2 residuals results at a few selected sites like Darwin.

This analysis was repeated with IRI's hmF2 model, showing that it overestimates for most sites across all seasons and solar cycles, but underestimates the Australasian sites during solar maximum. The standard deviations were fairly low, typically under $\sim$10 %. The lag analysis as per foF2 was performed with the hmF2 observations, this model was found to have similar lag problems.

Mocha provides moderate improvements to the foF2 magnitude model compared with IRI's and significant improvements to hmF2. Mocha reduces the lag in foF2, compared with reference data and reduces the RMS mean and RMS standard deviation of the foF2 and hmF2 residuals. Mocha also does not have the systematic overestimation of hmF2 as IRI. Mocha's tenth, fiftieth and ninetieth percentiles for foF2 and hmF2 all perform equally well, with low RMS means and RMS standard deviations. Mocha can be used as a replacement model for IRIs foF2 and hmF2 predictions. IRI may be run with user input foF2 and hmF2 which may be generated by Mocha.

Appendix A

This section describes the URSI descriptors and qualifiers used to filter the WDC data. Qualifiers indicate that the presented value has not been calculated directly. Lack of a qualifying letter indicates there were no difficulties in scaling the trace. Descriptive letters indicate the cause of the difficulty in the measurement. Each qualifier and descriptor is defined in Tables A.1 and A.2. A complete description is found within Piggott and Rawer (1978).

Letter	Qualifier	Descriptor
A	Less than. Used only with fbEs	Measurement influenced by, or impossible because of, the presence of a lower thin layer, for example, ES
B		Measurement influenced by, or impossible because of, absorption in the vicinity of fmin
C		Measurement influenced by, or impossible because of, any non-ionospheric reason.
D	Greater than	Measurement influenced by, or impossible because of, the upper limit of the frequency range in use.
E	Less than	Measurement influenced by, or impossible because of, the lower limit of the frequency range in use.
F		Measurement influenced by, or impossible because of, the presence of frequency spread.
G		Measurement influenced by, or impossible because the ionisation density of the reflecting layer is too small to enable it to be made accurately
H		Measurement influenced by, or impossible because of, the presence of stratification
I	Missing value has been replaced by an interpolated value	
J	Ordinary component characteristic deduced from the extraordinary component	
K		Particle E layer present.
L		Measurement influenced by, or impossible because the trace has no sufficiently definite cusp between layers. Mixed spread F present.
M	Interpretation of measurement questionable because ordinary and extraordinary components are not distinguishable. Used with descriptive letter which shows why components are not distinguishable	Interpretation of measurement questionable because ordinary and extraordinary components are not distinguishable. Used when interpretation is doubtful and a qualifying letter needed for other reasons (e.g., U, D, E)

Table A.1: the standard scaling parameters. The table is reconstructed from Piggott and Rawer (1978).

Letter	Qualifier	Descriptor
N		Conditions are such that the measurement cannot be interpreted.
O	Extraordinary component characteristic deduced from the ordinary component (Used for the x characteristic only.)	Measurement refers to the ordinary component
P		Man-made perturbations of the observed parameter; or spur type spread F present.
Q		Range spread present.
R		Measurement influenced by, or impossible because of, attenuation in the vicinity of the critical frequency
S		Measurement influenced by, or impossible because of, interference or atmospherics
T	Value determined by a sequence of observations, the actual observation being inconsistent or doubtful	Value determined by a sequence of observations, the actual value or observation being inconsistent or doubtful
U	Uncertain or doubtful numerical value	
V		Forked trace, which may influence the measurement
W		Measurement influenced by, or impossible because the echo lies outside the height range recorded
X		Measurement refers to the extraordinary component
Y		Lacuna phenomena, or severe F-layer tilt present
Z	Measurement deduced from the third magneto-electronic component	Third magneto-electronic component present

Table A.2: Describes the standard scaling parameters. The table is reconstructed from Piggott and Rawer (1978).

Appendix B

This appendix describes the effects of filtering on the number of World Data Centre (WDC) foF2 observations. There are two filtering types: filtering by the URSI parameters and storm filtering. URSI parameters are used to describe the conditions that were present when the measurement occurred. The URSI scaling parameters are described in detail in Appendix A. The storm filtering was done by filtering based of the Disturbance storm time (Dst) index. If a day had a Dst excursion below -75 nT then that day of observations were removed. After a moderate storm the strength of the ionosphere can take a day to recover, as such if a day had a Dst excursion below -100 nT then the subsequent day of observations was removed as well.

Figures B.1 and B.2 display the effects of the filtering level on Darwin and Vanimo observations respectively. The URSI filtering appears to affect Vanimo more significantly than Darwin, particularly during 2004 to 2008. The storm filtering also appears to affect Vanimo more than Darwin.

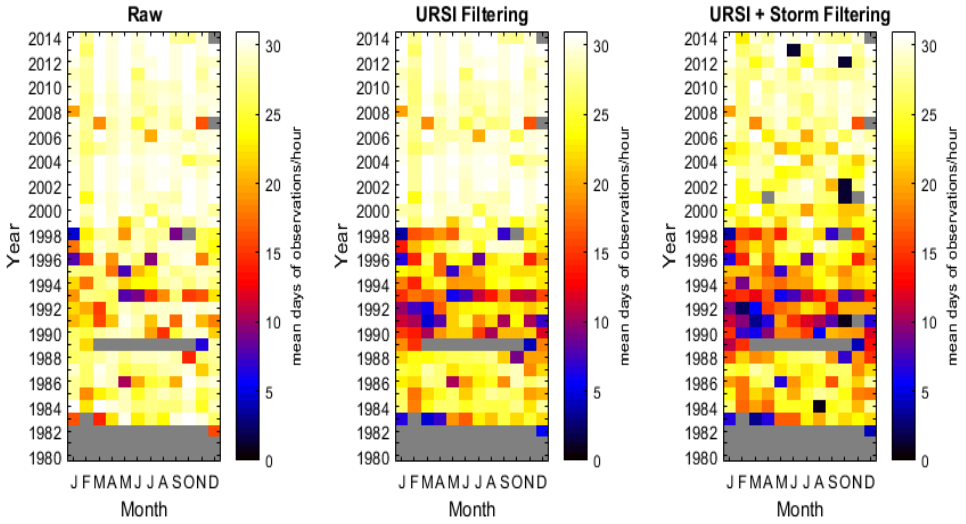

Figure B.1: Mean number of days of observations per hour of the raw WDC data (left) for Darwin, URSI parameter filtered (middle) and Storm and URSI filtering (right) between 1980 to 2014.

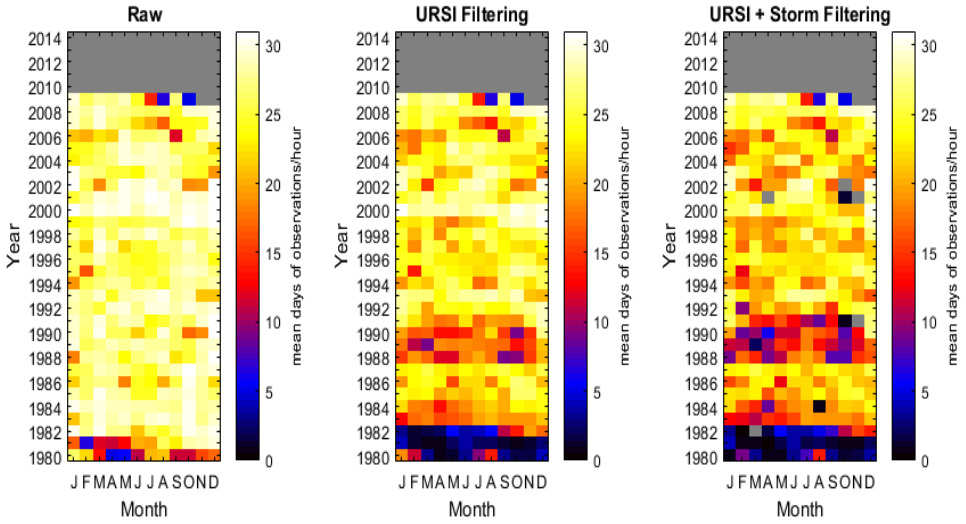

Figure B.2: Mean number of days of observations per hour of the raw WDC data (Left) for Vanimo, URSI parameter filtered (Middle) and the Storm and URSI filtering (Right) between 1980 to 2014.

Bibliography

Adebesin, B. O., Adekoya, B. J., Ikubanni, S. O., Adebiyi, S. J., Adebesin, O. A., Joshua, B. W. and Olonade, K. O. 2014, 'Ionospheric fof2 morphology and response of f2 layer height over jicamarca during different solar epochs and comparison with iri-2012 model', *Journal of Earth System Science*, vol. 123, no. 4, pp. 751–765.

Altadill, D., Magdaleno, S., Torta, J. and Blanch, E. 2013, 'Global empirical models of the density peak height and of the equivalent scale height for quiet conditions', *Advances in Space Research*, vol. 52, no. 10, pp. 1756 – 1769. IRI Over the African Region.

Bailey, G., Balan, N. and Su, Y. 1997, 'The sheffield university plasmasphere ionosphere model—a review', *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 59, no. 13, pp. 1541 – 1552. The Ninth International Symposium on Equatorial Aeronomy.

Bilitza, D. 1990, *International Reference Ionosphere 1990*, NSSDC 90-22, Greenbelt, Maryland.

Bilitza, D. 2001, 'International reference ionosphere 2000', *Radio Science*, vol. 36, no. 2, pp. 261–275.

Bilitza, D. 2004, '35 years of international reference ionosphere- karl rawer's legacy', *Advances in Radio Science*, vol. 2, pp. 283–287.

Bilitza, D., Altadill, D., Reinisch, B., Galkin, I., Shubin, V. and Truhlik, V. (2016), The International Reference Ionosphere: Model Update 2016, *in* 'EGU General Assembly Conference Abstracts', Vol. 18 of *EGU General Assembly Conference Abstracts*, p. 9671.

Bilitza, D., Altadill, D., Truhlik, V., Shubin, V., Galkin, I., Reinisch, B. and Huang, X. 2017, 'International reference ionosphere 2016 from ionospheric climate to real time weather predictions', *Space Weather*, vol. 15, no. 2, pp. 418–429.

Bilitza, D., Altadill, D., Zhang, Y., Mertens, C., Truhlik, V. et al. 2014, 'The international reference ionosphere 2012 – a model of international collaboration', *J. Space Weather Space Clim*, vol. 4. A07.

Bilitza, D., Brown, S. A., Wang, M. Y., Souza, J. R. and Roddy, P. A. 2012, 'Measurements and iri model predictions during the recent solar minimum', *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 86, pp. 99 – 106.

Bilitza, D., Eyfrig, R. and Sheikh, N. M. 1979, 'A global model for the height of the F2-peak using M3000 values from the CCIR numerical map', *ITU Telecommunication Journal*, vol. 46, pp. 549–553.

Bilitza, D., McKinnell, L.-A., Reinisch, B. and Fuller-Rowell, T. 2011, 'The international reference ionosphere today and in the future', *Journal of Geodesy*, vol. 85, no. 12, pp. 909–920.

Bilitza, D. and Reinisch, B. W. 2008, 'International reference ionosphere 2007: Improvements and new parameters', *Advances In Space Research*, vol. 42, pp. 599–609.

Clette, F. and Lefèvre, L. 2016, 'The new sunspot number: Assembling all corrections', *Solar Physics*, vol. 291, no. 9, pp. 2629–2651.

Clette, F., Svalgaard, L., Vaquero, J. M. and Cliver, E. W. 2014, 'Revisiting the sunspot number', *Space Science Reviews*, vol. 186, no. 1, pp. 35–103.

Davies, K. 1990, *Ionospheric Radio*, Electromagnetics and Radar Series, Peregrinus.

Dumbs, A., Emmenegger, G., Kist, R., Klumpar, D., Neske, E., Slavik, J., Spenner, K. and Wolf, H. 1979, 'Results from the plasma experiments on aeros', *Journal of geomagnetism and geoelectricity*, vol. 31, no. Supplement, pp. S125–S139.

Fabrizio, G. A. 2013, *High Frequency Over-The-Horizon Radar Fundamental Principles, Signal Processing, and Practical Applications*, McGraw-Hill, New York.

Galkin, I. A., Khmyrov, G. M., Kozlov, A. V., Reinisch, B. W., Huang, X. and Paznukhov, V. V. 2008, 'The artist 5', *AIP Conference Proceedings*, vol. 974, no. 1, pp. 150–159.

Harris, T. J. and Pederick, L. H. 2017, 'A robust automatic ionospheric o/x mode separation technique for vertical incidence sounders', *Radio Science*, vol. 52, no. 12, pp. 1534–1543.

Huba, J. D., Joyce, G. and Krall, J. 2008, 'Three-dimensional equatorial spread f modeling', *Geophysical Research Letters*, vol. 35, no. 10.

Ikubanni, S., Adeniyi, J. and Obrou, O. 2014, 'Monthly mean fof2 model for an african low-latitude station and comparison with iri', *Advances in Space Research*, vol. 53, no. 4, pp. 635 – 646.

IPS (1994), *HF Radio Propagation Course Manual*, Australian Government: Department of Industry Tourism and Resources, IPS Radio and Space Services Level 6, North Wig 477 Pitt Street, Sydney NSW 2000.

Jones, W. and Gallet, R. 1962, 'The representation of diurnal and geographic variations of ionospheric data by numerical methods', *Telecommun. J.*, vol. 29, no. 5, pp. 129–149.

Kaufmann, P. and Mendes, A. 1970, 'The correlation between sudden phase anomalies and solar microwave radio bursts', *Journal of Atmospheric and Terrestrial Physics*, vol. 32, no. 3, pp. 427 – 432.

Libo, L., Weixing, W. and Baiqi, N. 2003, 'Statistical modeling of ionospheric fof2 over wuhan', *Radio Science*, vol. 39, no. 2.

Mahajan, K., Aggarwal, S., Sethi, N. and Kohli, R. 1997, 'On the saturation of fof2 for r12 above 150', *Advances in Space Research*, vol. 20, no. 9, pp. 1761 – 1764. Quantitive Description of Ionospheric Storm Effects and Irregularities.

Mathworks (2016). The MathWorks, Natick, MA, USA, 01760-2098, USA.

McNamara, L. F. and Thompson, D. C. 2015, 'Validation of cosmic values of fof2 and m(3000)f2 using ground-based ionosondes', *Advances in Space Research*, vol. 55, no. 1, pp. 163 – 169.

Piggott, W. and Rawer, K. 1978, *U.R.S.I. Handbook of Ionogram Interpretation and Reduction*, second edn, World Data Centre A for Solar Terrestrial Phyiscs, NOAA, Boulder, Colorado.

Pignalberi, A., Pezzopane, M., Rizzi, R. and Galkin, I. 2018, 'Effective solar indices for ionospheric modeling: A review and a proposal for a real-time regional iri', *Surveys in Geophysics*, vol. 39, no. 1, pp. 125–167.

Pignalberi, A., Pezzopane, M., Tozzi, R., De Michelis, P. and Coco, I. 2016, 'Comparison between iri and preliminary swarm langmuir probe measurements during the st. patrick storm period', *Earth, Planets and Space*, vol. 68, no. 1, pp. 93.

Reinisch, B. W. and Galkin, I. A. 2011, 'Global ionospheric radio observatory (giro)', *Earth, Planets and Space*, vol. 63, no. 4, pp. 377–381.

Richmond, A. D., Ridley, E. C. and Roble, R. G. 1992, 'A thermosphere/ionosphere general circulation model with coupled electrodynamics', *Geophysical Research Letters*, vol. 19, no. 6, pp. 601–604.

Schunk, R. and Nagy, A. 2009, *Ionospheres, Physics, Plasma Physics and Chemistry*, Cambridge Atmospheric and Space Sciences, second edn, Cambridge University Press, The Edinburgh Building, Cambridge CB2 8RU, UK.

Schunk, R. W., Scherliess, L., Sojka, J. J., Thompson, D. C., Anderson, D. N., Codrescu, M., Minter, C., Fuller-Rowell, T. J., Heelis, R. A., Hairston, M. and Howe, B. M. 2004, 'Global assimilation of ionospheric measurements (gaim)', *Radio Science*, vol. 39, no. 1.

Shimazaki, T. 1955, 'World daily variability in the height of the maximum electron density of the ionospheric f2-layer', *J. Radio Res. Lab. (Japan)*, vol. , no. 2, pp. 85–97.

Shubin, V. 2015, 'Global median model of the f2-layer peak height based on ionospheric radio-occultation and ground-based digisonde observations', *Advances in Space Research*, vol. 56, no. 5, pp. 916 – 928.

Zhang, M.-L., Shi, J.-K., Wang, X., Shang, S.-P. and Wu, S.-Z. 2007, 'Ionospheric behavior of the f2 peak parameters fof2 and hmf2 at hainan and comparisons with iri model predictions', *Advances in Space Research*, vol. 39, no. 5, pp. 661 – 667.