Optimizing MIDAS III Over South Africa

A thesis submitted in partial fulfilment of the
requirements for the degree of

MASTER OF SCIENCE

of

RHODES UNIVERSITY

by

Nigussie Mezgebe Giday

September 2013

Abstract

In this thesis an ionospheric tomographic algorithm called Multi-Instrument Data Analysis System (MIDAS) is used to reconstruct electron density profiles using the Global Positioning System (GPS) data recorded from 53 GPS receivers over the South African region. MIDAS, developed by the Invert group at the University of Bath in the UK, is an inversion algorithm that produces a time dependent 3D image of the electron density of the ionosphere. GPS receivers record the time delay and phase advance of the trans-ionospheric GPS signals that traverse through the ionosphere from which the ionospheric parameter called Total Electron Content (TEC) can be computed. TEC, the line integral of the electron density along the satellite-receiver signal path, is ingested by ionospheric tomographic algorithms such as MIDAS to produce a time dependent 3D electron density profile. In order to validate electron density profiles from MIDAS, MIDAS derived NmF2 values were compared with ionosonde derived NmF2 values extracted from their respective 1D electron density profiles at 15 minute intervals for all four South African ionosonde stations (Grahamstown, Hermanus, Louisvale, and Madimbo). MIDAS 2D images of the electron density showed good diurnal and seasonal patterns; where a comparison of the 2D images at 12h00 UT for all the validation days exhibited maximum electron concentration during the autumn and summer and a minimum during the winter. A root mean square error (rmse) value as small as 0.88x $10^{11}[el/m^3]$ was calculated for the Louisvale ionosonde station during the winter season and a maximum rmse value of 1.92x $10^{11}[el/m^3]$ was obtained during the autumn season. The r^2 values were the least during the autumn and relatively large during summer and winter; similarly the rmse values were found to be a maximum during the autumn and a minimum during the winter indicating that MIDAS performs better during the winter than during the autumn and spring seasons. It is also observed that MIDAS performs better at Louisvale and Madimbo than at Grahamstown and Hermanus. In conclusion, the MIDAS reconstruction has showed good agreement with the ionosonde measurements; therefore, MIDAS can be considered a useful tool to study the ionosphere over the South African region.

Acknowledgements

Firstly, I would like to express my appreciation to my supervisors, Dr Zama Katamzi and Dr Lee-Anne McKinnell, for their invaluable support and contribution to this research.

My gratitude goes to the National Astrophysics and Space Science Program (NASSP) for their financial support throughout the programme. I would also like to thank the South African National Space Agency (SANSA) Space Science Direcorate for the use of their academic facilities and residential service.

MIDAS is a property of the Invert group at the University of Bath, so I would like to give recognition for their input. I would also like to acknowledge the Chief Directorate Surveys and Mapping (CDSM) for allowing the GPS data used in this study.

The support I got from Dr Ben Opperman in making use of the GPS data in MIDAS algorithm and Dr Pierre Cilliers for providing me with explanations that supplement my thesis is greatly appreciated. I would also like to thank Dr John Bosco Habarulema for his support in latex techniques, Mr. Nicholas Ssessanga and Mr. Vumile Tyalimpi for their support in MATLAB programming techniques.

My heartfelt appreciation goes to my colleagues (Thabang Matladi, Thinawanga Tshilande, Lifa Mbuli, Tshimangadzo Matamba, Ahoua Malan, Electdom Matandirotya, Minko Flavien, Sarina Molepo, Sfundo, Arios, Bafana and others) for making my stay at SANSA Space Science both very joyful and memorable. The time we spent together including the Soccer games will be memorable. I also thank the SANSA Space Science staff for their support and kindness throughout the study.

Last but not least, my sincere thanks and gratitude go to my parents and my uncle Mr. Tsegay Redda for supporting me both in material and most importantly for allowing me to follow 'My Journey in life'.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

The research presented in this thesis aims to optimize an ionospheric tomography system called Multi-Instrument Data Analysis System (MIDAS) over the South African region. MIDAS is an inversion technique developed by the invert group at the University of Bath in the United Kingdom (UK). A feasibility study of the first MIDAS version (MIDAS1) was conducted by Katamzi (2008). The study was to reconstruct one-dimensional electron density profiles using Global Positioning System (GPS) data from 8 to 10 GPS receivers over the South African region and verify it with ionosonde derived one-dimensional electron density profiles at the Grahamstown South Africa ionosonde station.

In this study, the third version of MIDAS (MIDAS3) is used to reconstruct a time varying three-dimensional ionospheric electron density profile using GPS data from around 53 GPS receivers over the South African region. One dimensional electron density profiles derived at each of the four ionosonde stations available in South Africa, were used to validate similar dimension profiles reconstructed from MIDAS3. For convenience MIDAS3 will be referred to as MIDAS for the rest of this thesis.

The ionosphere is an ionized medium which refracts, delays and reflects electromagnetic waves that traverse through the medium. High frequency (HF) communication

is made possible due to the reflective nature of the ionosphere (McNamara, 1991); therefore, it is important to study the characteristic features and dynamics of the ionosphere so that a near real-time variation of the ionosphere can be traced and mapped using ionospheric measurement tools. In fact, by having an organized near real-time 3-D global map of electron density updated regularly, the radio frequency system user can apply corrections both where and when needed (Bust and Mitchell, 2008). Using ionospheric tomography systems such as MIDAS, time varying three-dimensional electron density profiles of the ionosphere can be mapped (Mitchell and Spencer, 2003). An introduction to the ionosphere, the layers of ionosphere and its effect on HF signals will be discussed in this chapter. The objectives of this thesis and a general overview of the remaining chapters will also be discussed within this chapter.

1.1 Research Aim

The research described within this thesis is aimed at optimizing an ionospheric tomography system, MIDAS, which ingests data from GPS satellites to reconstruct three dimensional and time varying electron density within a defined ionospheric region (Mitchell and Spencer, 2003). The project main aim can be divided into two parts, namely:

- To optimize ionospheric tomography imaging over South Africa using the extensive South African GPS network so as to derive a reasonable electron density profile.

- To verify the derived electron density profile with independent measurements, such as from the four South African ionosondes.

1.2 General Overview of the Ionosphere

1.2.1 The Ionosphere

The ionosphere, a region of ionized plasma, results from an interaction of solar emissions with the terrestrial atmosphere. The bottomside ionosphere begins near 50 km in altitude and extends to roughly 1000 km above the Earth's surface where it merges with the magnetosphere (Selcher, 2007). This region of the atmosphere exhibits different characteristics during day and night times. As the sun rises ultra-violet (UV) radiation starts to break up gas molecules into positively charged ions and free electrons. The electron density builds up, with a peak at around 2 p.m local time, and then starts declining (Misra and Enge, 2006; Rose, 2011). At night the process of ionization ends and the ions and free electrons recombine, reducing the number of free electrons by one to two orders of magnitude (Misra and Enge, 2006, page 162). Since the source of ionization is the Sun's radiation, the thickness and nature of dominant chemical processes within the ionosphere vary between day and night; moreover, the ionosphere changes with seasons and phases of the eleven-year solar cycle (McNamara, 1991, page 22). The evolution of the ionospheric region is influenced by changes in solar emissions, atmospheric dynamics and the interplanetary and geomagnetic fields (Mitchell and Spencer, 2003).

The bottomside ionosphere can be divided into four layers which vary in altitude and depend on the rates of production and loss of free electrons at each altitude.

1.2.2 Layers of the Ionosphere

D-Layer:

This region is the lowest layer of the ionosphere extending from approximately 50 km to 90 km above the Earth's surface. The ionization of atomic nitrogen to produce

nitric oxide (NO) is the characteristic feature of the D-region. The D-layer is a highly collisional, very weakly ionized region with primary components of molecular oxygen ions (O_2^+) and nitric oxide ion (NO^+) (Selcher, 2007). The typical electron density of the D-layer is of the order of 10^3 el/cm^3. Sources of ionization in the D-region are X-rays, cosmic rays and solar Lyman α. The solar Lyman α ionizes only the dominant species NO. Details of the chemical composition of this layer can be found in Rishbeth and Garriott (1969). The D-layer is mostly seen during the day and disappears during night time; however, it sometimes remains to appear during the night time due to the high energy galactic cosmic rays that reach this layer. The intensity of these ionizing sources reaching these heights is small; resulting in low electron density concentration. Hence, the contribution of the D-layer to the ionospheric total electron content (TEC) is negligible. Therefore, it has no significant role in ionospheric tomography. However, this layer is responsible for the absorption of low frequency radio waves.

E-Layer:

The E-layer is immediately above the D-layer, and extends approximately from 90 km to 120 km. Soft X-rays absorbed by molecules in this region are the source of ionization. There is a relatively higher concentration of electron density in this region than in the D-layer. This layer only reflects radio waves with a frequency lower than 10 MHz and partially absorbs radio waves of higher frequencies (McNamara, 1991). Typical E-region peak electron density is of the order of 10^5 el/cm^3 (Selcher, 2007). At mid-latitudes, sporadic E (Es) layers are made up of clouds of electrons a few kilometers thick and a few hundred kilometers across, at an altitude approximately between 90 km and 120 km (McNamara, 1991). Sporadic E is a reflecting layer which appears and disappears sporadically at the E-region altitudes of the ionosphere. At mid-latitudes, Es is mostly a daytime summer phenomenon. Midday values of the foEs (critical frequency of the Es layer) reach around 10 MHz during summer.

Statistical occurrences of foEs values greater than 5 MHz in mid-latitude regions can be found in McNamara (1991).

F-Layer:

The F-layer extends from an altitude of about 120 km to around 400 km. Photoionization of atomic oxygen by extreme EUV radiation from the sun is the source of ionization in this region. The layer splits into two layers, F_1 and F_2, during daylight and the former disappears to form one F-layer during the night time. This layer is the most important of all the ionospheric layers in that it makes communication using HF waves over the horizon possible (Baumjohann and Treumann, 1997; McNamara, 1991, page 19). This layer has the most dense free electrons compared to the other layers; therefore, is the most contributor to ionospheric TEC.

The ionosphere is most often divided into two main sections, the bottomside ionosphere consisting of the D-, E- and F- layers, and the topside ionosphere which extends above the peak of the F-layer. The distribution of different species of molecules at different layers of the ionosphere is mainly due to the difference in molecule's density. Lighter species of molecules reside at higher altitudes pushing the heavier molecules towards lower altitudes. The lower and bottomside ionosphere regions contain more ionizable neutral molecules. The intensity of the ionizing solar radiation that reaches the upper region of the ionosphere is high; however, as the ionizing solar radiation penetrates down it encounters an increasing number of ionizable neutral molecules which reduces its intensity significantly. Different wavelengths of the solar radiation are absorbed at different altitudes of the ionosphere causing a number of chemical reactions which results in numerous positively and negatively charged species to be dominant at different layers of the ionosphere. For more information see Rishbeth and Garriott (1969).

1.2.3 Topside Ionosphere

The topside side ionosphere extends from the F_2 peak to around 1000 km and is populated by a mixture of ions (O^+, He^+, H^+) (Bauer, 1962; Sibanda, 2010). At higher altitudes the H^+ ion is predominant. Sibanda (2010) pointed out that the electron density distribution of the topside ionosphere is closely linked to the dynamics of the lower bottomside ionosphere. In his work Sibanda (2010) also pointed out that the F-region is considered a source and drain of ionization for the topside ionosphere. Figure 1.1 illustrates the ionospheric layers and the distribution of ionospheric dominant species with altitude.

1.2.4 Variability of the Ionosphere

The ionosphere is known to vary with changes in hour, season, location, and solar cycle.

Diurnal Variability

The variation in solar illumination of the ionosphere is the main cause of the diurnal variation of the ionosphere. However, at high latitudes the $\mathbf{E} \times \mathbf{B}$ convection of ionization from daylit latitudes toward the night side is a factor for diurnal variation of the ionosphere. Evidence of this occurrence is the existence of the F_2 layer during the long winter night at these latitudes (Davies, 1990, pages 124-125). Large diurnal variations occur particularly in the lower layers of the ionosphere (D- and E-layers). The intensity of solar radiation increases as the sun rises and decreases as the sun sets; hence as the sun sets the ionization process falls causing the D- and the E-layers to disappear at night time. The F_1 and F_2 layers combine to form one F- layer whose peak height rises. The electron densities are greater at all heights during day than at night.

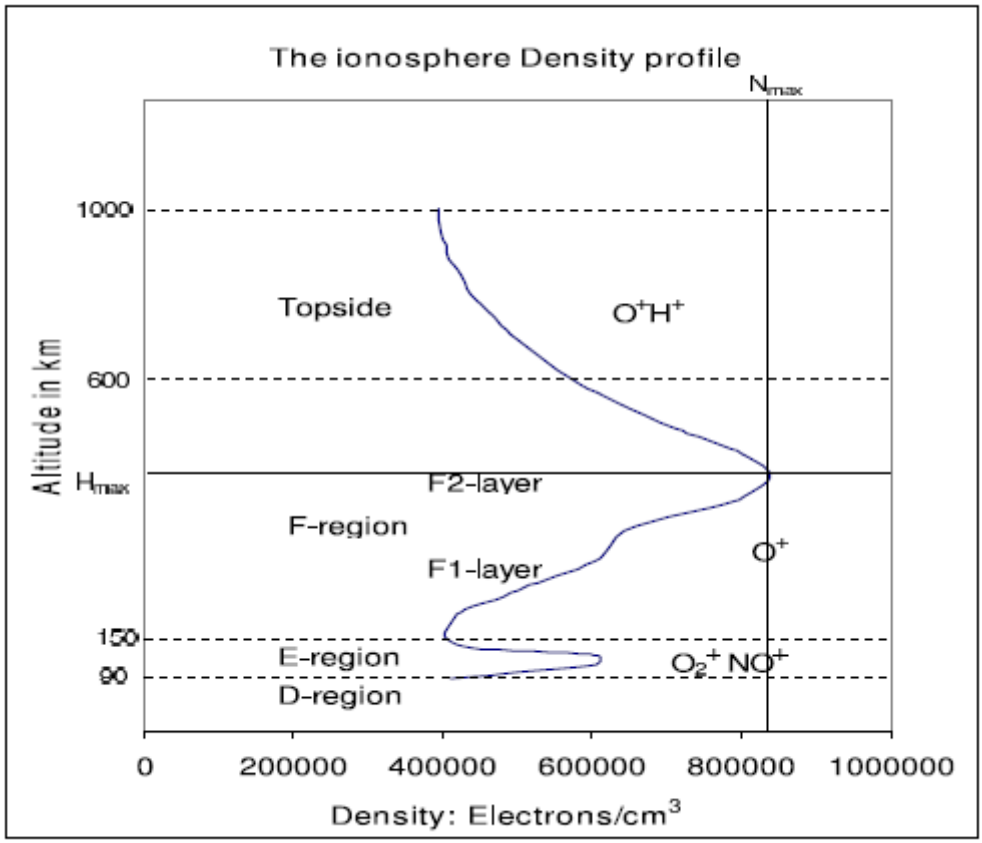

Figure 1.1: An illustration of ionospheric layers along with dominant species in each region. *after* Katamzi (2008).

Seasonal Variability

The ionosphere also exhibits seasonal variability mainly in the variation of the peaks of the electron density concentration and the falling and rising of peak electron density heights. Seasonal variations are the result of the Earth revolving around the Sun; the relative position of the Sun moves from one hemisphere to the other with changes in seasons. The seasonal variations of the layers of the ionosphere correspond to the variation in the highest zenith angle of the sun; summer being the season with high ionization density in the D-, E- and F_1 layers (Rishbeth and

Garriott, 1969, page 179). At lower latitudes, the night time F-layer tends to be at higher altitudes in summer than in winter (Davies, 1990, page 126). The night time F-layer tends to be thicker when higher. In general, the maximum electron density and the total electron content of the nighttime F-layer are greater in summer than in winter. When it comes to middle latitudes the peak density is considerably larger in winter than in summer at noon; a phenomenon called the mid-latitude seasonal anomaly. Davies (1990) also showed that at the middle latitudes the peak density decreases at a relatively higher altitude during the summer months.

Variability with Location

The Earth's ionosphere also varies with geomagnetic latitude in that the equatorial ionosphere possesses a thicker F_2 layer than elsewhere; moreover, there is a region of high electron concentration at geomagnetic latitudes of about $\pm 20^{\circ}$ during the early evening (Davies, 1990; McNamara, 1991; Sibanda, 2010). This phenomenon is driven by the fountain effect: an electromagnetic uplifting of the plasma until the pressure and the gravity forces become high enough to force the plasma back through the magnetic field lines to higher latitudes. Sibanda (2010) also pointed out that the middle latitude ionosphere shows least variation compared to the high and low latitude regions. According to Yizengaw (2004) seasonal anomalies and short term variations like the main ionospheric trough (MIT) driven by solar and geomagnetic storm inputs in the auroral ionosphere occur in these regions. The high latitude ionosphere is very variable both in space and time (McNamara, 1991, page 98) because particles from the sun carried by the solar wind can penetrate into the low levels of the ionosphere and collide with the neutral atmosphere. This can cause an increased ionization in a process known as collisional ionization. And this becomes a source of problem for HF communications in this region.

Variation with Solar Activity

The sunspot number is an indicator of the solar activity. Ionization of the ionosphere is driven by solar electromagnetic and corpuscular radiation, therefore it is a function of solar activity. The variation in solar activity oscillates between periods of solar maximum (maximum number of sunspots) and solar minimum (minimum number of sunspots). The period of solar activity is approximately 22 years with an approximate 11 year sunspot cycle. When solar activity reaches its peak the intensity of the radiation that reaches the Earth's atmosphere increases thus enhancing the electron concentration in the ionosphere. During periods of maximum solar activity, a relatively large increase in electron density occurs at successively higher altitudes. At these times, higher operating frequencies must be used for long distance communications since the critical frequencies (maximum frequencies associated with peak electron densities of the ionospheric layers) of the ionospheric layers are higher (McNamara, 1991; Sibanda, 2010).

Magnetospheric disturbances driven by the Sun's activities are also major causes of the ionosphere's variation. The magnetosphere is the region above the Earth's ionosphere in which the magnetic field of the Earth has a dominant control over the motions of gas and fast charged particles (Van Allen, 2004, page 72). During geomagnetic storms the disturbed solar wind compresses the Earth's magnetosphere. During the interaction, intense electric fields occur and are mapped along the geomagnetic field lines to the higher latitude ionosphere (Yizengaw, 2004). When these penetrate into lower altitudes, the precipitated energetic particles increase ionospheric ionization at higher latitudes (Yizengaw et al., 2005).

1.2.5 Plasma Frequency

Electrons in the ionosphere oscillate with a characteristic frequency known as the plasma frequency. Chen (1984) showed that the characteristic plasma frequency at any point in the plasma is given by:

$$\omega_p = 2\pi f = \sqrt{\frac{ne^2}{\varepsilon_0 m}} \tag{1.1}$$

where

f is the corresponding plasma frequency (in Hertz)

n is the number density of the electrons (in electrons per cubic meter)

$e = 1.60 \times 10^{-19}$ C is the electron charge

$\varepsilon_0 = 8.85 \times 10^{-12}$ Fm^{-1} is the permittivity of free space, and

$m = 9.11 \times 10^{-31}$ kg is the electron mass

Substituting the values for the constant parameters in Equation 1.1 and normalizing the units of the frequency to MHz, the equation for electron density (in units of el/m^3) is reduced to a more general and simpler form of:

$$n = 1.24 \times 10^{10} f^2 \tag{1.2}$$

Equation 1.2 shows a square-law relation between the characteristic plasma frequency of the ionosphere, f, and the corresponding electron density, n.

The plasma frequency is the highest radio frequency that a portion of the ionosphere is able to reflect if radio waves are at normal incidence to the ionosphere. A radio signal penetrates or traverses a portion of the ionosphere if the radio frequency is greater than the plasma frequency. If the radio frequency is less than or equal to the plasma frequency, the signal is reflected by the ionosphere (Okoh, 2009). The ionospheric interaction with the signals depends on the amount of electron density.

For this reason, electron density is one of the most important parameters to look at when studying the state of ionospheric activities.

1.3 Total Electron Content Parameter

The ionosphere affects modern technologies such as civilian and military communications, navigation systems, surveillance systems *etc* mainly because they all use HF waves that traverse the ionosphere. However, the GPS and HF radio signals traveling through the ionosphere are refracted, reflected, dispersed and subjected to time delay and phase advance. For many communication and navigation systems this is an issue because the systems use signals transmitted to and from satellites which must pass through the ionosphere. In order to get reliable communication and navigation services it is necessary to correct the signals from errors that are introduced by the ionosphere. To do that, the properties of the ionospheric parameters and their variability with respect to magnetospheric disturbance, time of day, season of the year and solar cycle variability must be well understood. Apart from its use to draw correction measures to radio frequencies, characterizing the ionosphere is also vital for space weather effects and communication services. An important measurement of the ionosphere is the total electron content (TEC) which is an integrated electron density along the ionospheric ray path between the satellite and the receiver. Ionospheric imaging uses TEC measurements to produce two-, three-, and four-dimensional maps of electron density (Bust and Mitchell, 2008).

TEC can be measured by a number of standard techniques such as Faraday rotation, group delay and dispersive carrier phase (Yizengaw, 2004). The TEC estimation technique from GPS satellite signal group delay and dispersive carrier phase advance is a popular method. This is because studying the ionosphere using TEC measurements covers the total range of ionosphere. The time delay the radio waves

experience is proportional to TEC which is inversely proportional to the square of the signal frequency (Misra and Enge, 2006, page 164). However, ground and space based data sets such as the incoherent scatter radars (ISRs), ionosondes and satellite in-situ measurements are also important ionospheric measurements (Bust and Mitchell, 2006). For this study TEC values obtained from GPS satellites will be used as an input for a process called ionospheric tomographic inversion to determine the electron density profiles of the ionosphere. In this thesis MIDAS, one of the many ionospheric tomographic methods, is used to reconstruct electron density profiles of the ionosphere over the South African region.

1.4 Overview of the Thesis

The thesis is presented in 6 chapters including the introduction as Chapter 1. Chapter 2 discusses the details of the available data sources used in this research, the methods used to obtain TEC values from GPS signals and electron density profiles of ionosonde data.

In chapter 3 a discussion of ionospheric tomographic inversion is presented. Theoretical background of what ionospheric tomography is and a literature review of different tomographic algorithms is discussed. The application of empirical models and/or functions to assist tomographic inversion algorithms in regions where there is sparsity in data is also addressed in chapter 3. In this study, the Chapman function and the International Reference Ionosphere (IRI) models are discussed.

MIDAS, the inversion algorithm used in this study is discussed in chapter 4. The basics of the algorithm, the inversion methods involved and the procedures followed to obtain the results from the inversion process are presented in this chapter.

MIDAS outputs time varying three-dimensional electron density profiles. The 2D

images of electron densities will be illustrated and discussed in chapter 5. In addition, NmF2 values extracted from one dimensional electron density profiles reconstructed using MIDAS will be compared with NmF2 values extracted from similar dimension profiles from ionosonde measurements. The coefficient of determination (r^2) and root mean square errors are used to compare the peak electron density values reconstructed from MIDAS with peak electron densities from ionosonde measurements. The discussion for evaluation methods of the reconstructed and measured profiles is included in this chapter.

Chapter 6 presents a summary and discussion of the results of the research presented in this thesis, and a brief discussion about future work direction.

Chapter 2

Available Data Sources and Measurements

2.1 Global Positioning System

The Global Positioning System (GPS) is an all-weather, worldwide, continuous coverage, satellite-based radio navigation system (NAVSTARGPS, 1996). GPS is part of the Global Navigation Satellite System (GNSS) and was developed by the United States Department of Defense (DoD) for instantaneous determination of position and velocity (Hofmann-Wellenhof et al., 1997). The GPS constellation of satellites is revolutionizing the science and technology of the Earth's ionosphere (Mannucci et al., 1999). This is because, GPS is a unique resource of ionospheric measurements since it provides instantaneous global coverage, high resolution and near real-time data acquisition. It was initially introduced to overcome the limitations that the US military force was facing in navigation systems (Christie et al., 1996); however, through continuous advancement and development the system became a source of information for both the military and civilians. Currently GPS provides a source of information for conducting scientific studies as well as for navigation systems and land surveying, to mention a few.

GPS consists of three major segments: space, control, and user (NAVSTARGPS, 1996). The space segment consists of the GPS satellites. Only 24 of the 31 satellites in orbit are operational and the remaining are there as backup. The satellite vehicles (SVs) orbit the earth in six 12-hour (11-hour 58-minute) orbital planes with four satellites in each plane. The orbits are nearly circular with inclination angles of 55° relative to the equator and radii of approximately 20 200 km from the surface of the Earth. The six orbits are equally spaced around the equator with 60° separation. Since each satellite makes a complete orbit in almost 12 hours, each one of them passes over the same location on Earth each day. This constellation ensures that a user located anywhere on the globe is in direct line of sight to at least 4 satellites at any given time. GPS satellites emit coded radio signals continuously on two separate frequencies, 1.575 GHz and 1.227 GHz, referred to as L1 and L2 band respectively. The satellites can also transmit additional RF signals at frequencies referred to as L3 and L4 associated with classified onboard payloads. L3 is associated with Nuclear Detonation Detection System (NDS) and L4 with a payload called Reserve Auxiliary Package (RAP) (Misra and Enge, 2006). For further reading see Farrell and Barth (1998) and NAVSTARGPS (1996). Figure 2.1 shows a constellation of GPS satellites.

The other segment of the GPS system is the ground-based control system. This segment is responsible for monitoring the health and status of the space segment. The functions of the control units range from tracking the satellite trajectory, monitoring telemetry, updating the satellite ephemeris data (contains information to calculate position of each satellite) and almanac (information about the time and status of the entire satellite constellation) to disseminating useful information for the user segment. There are five unmanned monitor stations and one Master Control Station (their geographical locations are given in NAVSTARGPS (1996), page 15). The five unmanned stations monitor GPS satellite signals and then send that information

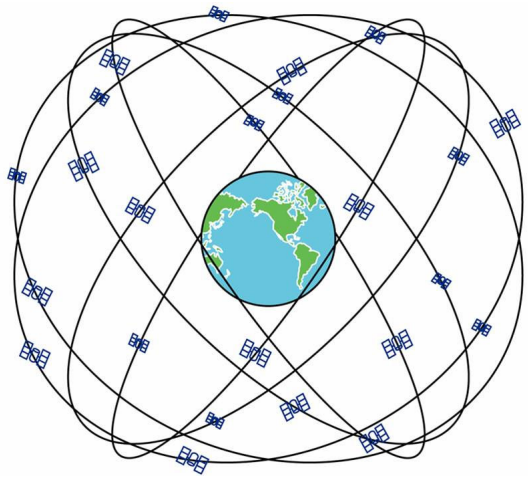

Figure 2.1: The figure illustrates the Constellation of GPS satellites taken from a PhD thesis by Julian A. Rose, University of Bath, 2011.

to the Master Control Station where anomalies are corrected and sent back to the GPS satellites through ground antennas. Detailed discussion can be found in Misra and Enge (2006), pages 34-35 and NAVSTARGPS (1996).

The user segment consists of antennas and receiver-processors that measure and decode the satellite transmissions to provide positioning, velocity, and precise timing information to the user. The space segment can provide service to limitless simultaneous users since user receivers operate passively. Receivers can be stand-alone, integrated with or embedded into other systems (NAVSTARGPS, 1996).

The ground GPS receivers are available in either single or dual frequency receivers. A single frequency receiver receives signals on L1 band frequency only and suffer from the effect of ionospheric delay while a dual frequency GPS receiver receives both the L1 and L2 band frequency signals simultaneously which enables it to reduce ionospheric effects. In fact, the dual frequencies (L1 and L2) of the dual frequency receivers are essential for eliminating the major part of ionospheric refraction

(Hofmann-Wellenhof et al., 1997; Hoque and Jakowski, 2012). Therefore, dual frequency GPS receivers provide more accurate results than the single frequency GPS receivers. The L1 frequency contains the civilian Coarse Acquisition (C/A) Code as well as the military Precise (P) Code. The L2 band frequency contains only the P code which is encrypted by the military using a technique called Anti-Spoofing (AS). Anti-spoofing is a technique whereby the P code is encrypted to the Y mode (which requires authorization) so that the military can obtain Precise Positioning Service (PPS) information. The single frequency GPS receivers are only able to provide Standard Positioning Service (SPS), which is based on C/A code and is, therefore, less precise than the dual frequency GPS receivers. For further information see Farrell and Barth (1998), pages 142-144 and Misra and Enge (2006), pages 37-39.

2.1.1 GPS Signal Propagation through the Ionosphere

Radio waves are delayed, refracted and dispersed by the ionosphere as they travel through it (Rose, 2011). In addition, GPS signals suffer from dispersion owing to the frequency dependence of the ionosphere's refractive index. To have an understanding of the time delays and phase advances of the GPS signals experienced due to the refractive nature of the ionosphere, it is important to discuss magneto-ionic theory with due attention to propagation of the signal radio waves through the ionosphere.

The frequency dependent refractive index of the ionosphere is given by the Appleton-Hartree formula (Davies, 1990).

$$\mu^2 = 1 - \frac{X}{1 - iZ - \frac{{Y_T}^2}{2(1 - X - iZ)} \pm \left[\frac{{Y_T}^4}{4(1 - X - iZ)^2} + Y_L^2 \right]^{\frac{1}{2}}} \tag{2.1}$$

where

$$X = \frac{\omega_{pe}^2}{\omega^2}$$

$$Y = \frac{\omega_{ce}}{\omega}$$

$$Y_L = Y cos\theta$$

$$Y_T = Y sin\theta$$

$$\omega_{pe}^2 = [2\pi f_{pe}]^2$$

$$\omega_{ce}^2 = [2\pi f]^2 = \frac{B_e}{m_e}$$

$$Z = \nu/\omega$$

in which μ stands for the ionospheric refractive index; ω_{pe} is the angular electron plasma frequency; ω is the angular frequency of the propagating radio wave; ω_{ce} is the angular gyro-frequency of the electrons; f_{pe} is the electrons plasma frequency; f is the radio wave frequency; Y_T and Y_L refer to the transverse and longitudinal components of the imposed magnetic fields with respect to the direction of the waves propagation and θ is the angle between the Earth's magnetic field and the waves propagation path. The ordinary and extra-ordinary wave modes are represented by the plus and minus signs respectively.

For the propagation of high-frequency radio waves through the E- and F- regions of the ionosphere, the collision term Z is very small and can be neglected (Rose, 2011). This is because collisions between electrons and neutral atoms is negligible due to low density of neutral atoms in these regions. With this assumption and some mathematical simplifications, Equation 2.1 can be rewritten as:

$$\mu^2 = 1 - \frac{X(1-X)}{1 - X - \frac{{Y_T}^2}{2} \pm \left[\frac{{Y_T}^4}{4} + (1-X)^2 Y_L^2\right]^{\frac{1}{2}}} \tag{2.2}$$

With the assumption that the magnetic field is negligible, whereby Y = 0, Equation 2.2 further reduces to:

$$\mu^2 = 1 - X \tag{2.3}$$

At GPS frequencies, the higher order terms can be neglected and the refractive index can be approximated (Parkinson and Spilker, 1996) by:

$$\mu \approx 1 - \frac{1}{2}X \tag{2.4}$$

in which

$$X = \frac{\omega_{pe}^2}{\omega^2} = \frac{[2\pi f_{pe}]^2}{[2\pi f]^2} \tag{2.5}$$

and

$$f_{pe} = \frac{ne^2}{4\pi^2 m_e \varepsilon_0} \tag{2.6}$$

Substitution of Equation 2.5 into 2.4 along with the constants gives the refractive index of the medium as:

$$\mu = 1 - \frac{40.31}{f^2}n \tag{2.7}$$

where 40.31 is a constant with units of $m^3 s^{-2}$.

The relationship between phase and group refractive indices is given by (Hofmann-Wellenhof et al., 1997):

$$\mu_g = \mu_p + f\frac{d\mu_p}{df} \tag{2.8}$$

which allows the phase and group refractive indices of the medium to be respectively written as:

$$\begin{aligned} \mu_p &= 1 - \frac{40.31}{f^2} n \\ \mu_g &= 1 + \frac{40.31}{f^2} n \end{aligned} \tag{2.9}$$

The fact that $\mu_g > \mu_p$ and, $v_g < v_p$ (since $v = \frac{c}{\mu}$) implies that GPS code measurements are delayed and the carrier phases advanced which in turn infers that the code psuedoranges are measured too long while the carrier phase pseudoranges are measured too short relative to the actual satellite and receiver geometric range. The difference in both measurements is the same (Hofmann-Wellenhof et al., 1997). It can be seen from Equation 2.7 that the refractive index is slightly less than one, and the phase velocity of the GPS carriers in the ionosphere exceeds that of light in a vacuum by enough to be significant for precise positioning (Misra and Enge, 2006, page 163).

2.1.2 Phase Advance and Group Delay of GPS Signals

The measured range between satellite S and receiver R along the signal path, according to Fermat's principle (see Hofmann-Wellenhof et al. (1997), page 102) is:

$$\rho = \int_R^S \mu d\rho. \tag{2.10}$$

Geometric range, the distance that could have been measured if the ionospheric refractive index, $\mu = 1$, is:

$$S = \int_R^S 1 d\ell. \tag{2.11}$$

The difference between measured and geometric range called ionospheric refraction (Hofmann-Wellenhof et al., 1997, page 102) is given by:

$$\Delta\rho = \rho - S = \int_R^S \mu d\rho - \int_R^S 1 d\ell. \tag{2.12}$$

Substituting the value for μ from Equation 2.9 and assuming the integration in Equation 2.10 to be along the geometric path, thus setting dρ= dℓ, the group delay becomes:

$$\Delta\rho_g = \frac{40.31}{f^2} \int_R^S n d\ell \tag{2.13}$$

A similar procedure for the phase advance gives:

$$\Delta\rho_p = -\frac{40.31}{f^2} \int_R^S n d\ell \tag{2.14}$$

Total Electron Content (TEC) is defined as the line integral of the electron concentration along the signal path from a satellite, S, to a receiver, R (Mitchell and Spencer, 2003). Mathematically, it can be written as:

$$TEC = \int_R^S N_e(r, \theta, \phi, t) ds \tag{2.15}$$

in which

N_e is the electron concentration

r is the radial distance from the center of the Earth

θ, ϕ are the latitude and longitude respectively

t is the time, and s is the distance along the satellite-to-receiver ray path.

Taking the integral part, in Equations 2.13 and 2.14, to be the definition of Total Electron Content (TEC), the equations for the carrier phase advance and group delay may be rewritten as:

$$\Delta\rho_g = \frac{40.31}{f^2} TEC \tag{2.16}$$

$$\Delta\rho_p = -\frac{40.31}{f^2}TEC \tag{2.17}$$

TEC measurements can be measured in slant TECs (sTEC), TEC measurements along a slant path, and vTEC, TEC measurements along the vertical direction (Chartier et al., 2012). A TEC unit (TECU) is the unit of TEC which is equivalent to 10^{16} electrons per square meter. Figure 2.2 shows the pictorial representation of sTEC and vTEC .

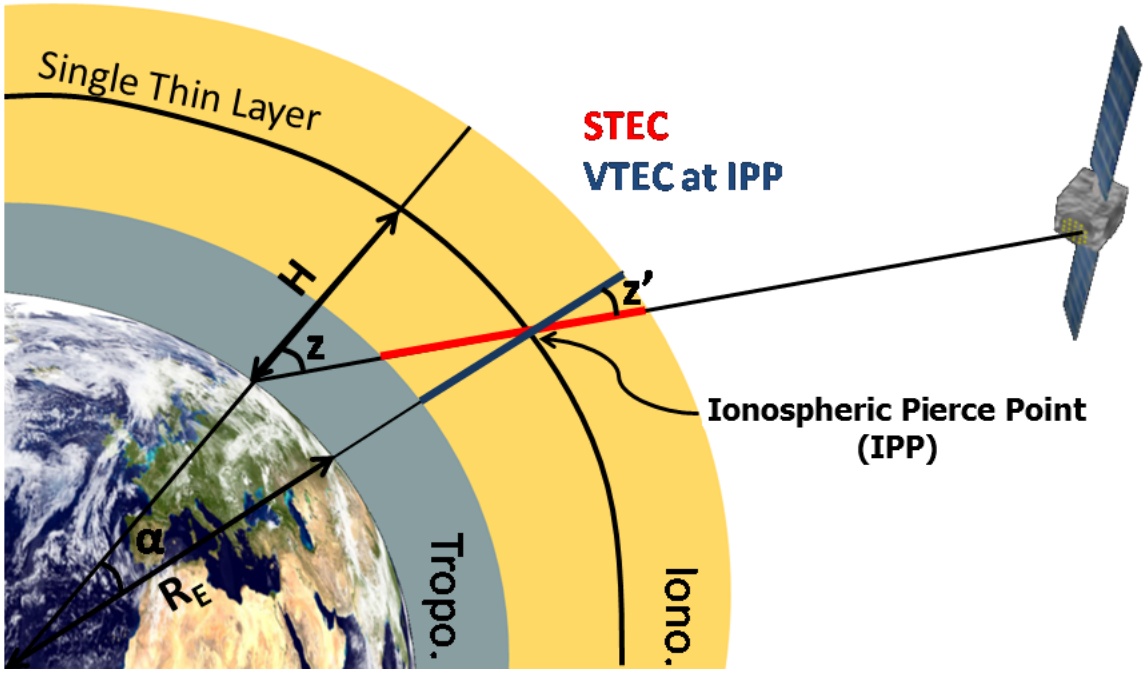

Figure 2.2: Figure showing the definition of sTEC and vTEC. Source *http : //gnss.be/ionosphere_tutorial.php*

2.1.3 TEC Measurements Using GPS

A wealth of information about the ionosphere can be recovered from GPS TEC measurements. However, in addition to the errors introduced by the ionosphere, satellite and receiver biases, satellite clock error and multi-path errors (discussed in Allain and Mitchell (2009) and Rose (2011)) make the direct measurement of TEC difficult. Hence, direct measurements from the differential code delay and phase advance measurements on L1 and L2 frequencies is used to overcome the problem. Since the C/A code (C1) on L1 frequency is unencrypted, its corresponding P code (P1) can be used by all GPS receivers. Similarly, the encrypted P code (P2) on L2 frequency can be used by tracking it with the P1 code (Allain and Mitchell, 2009). Ground GPS receivers record L1 and L2 data in RINEX (Receiver INdependent Exchange) format. RINEX format is the current standard format for GPS receivers (Misra and Enge, 2006, page 553). RINEX files are text files that list the L1 and L2 carrier phase and pseudocode path lengths (observables) to each satellite in view, sampled every 1 or 30 second intervals (Allain and Mitchell, 2009; Meggs et al., 2006). The observables are travel time or ranges which are computed from measured time or phase differences based on a comparison between received signals and receiver generated signals (Hoque and Jakowski, 2012). Therefore, the ranges are biased by satellite and receiver clock errors, instrumental biases and atmospheric effects (Rose, 2011; Allain and Mitchell, 2009), hence, called psuedoranges. The RINEX files of GPS receivers contain the following observables (Rose, 2011, page 35).

$$P_1 = P_0 + \frac{TEC}{f_1^2} + \varepsilon_1 \tag{2.18}$$

$$P_2 = P_0 + \frac{TEC}{f_2^2} + \varepsilon_2 \tag{2.19}$$

$$L_1 = P_0 - \frac{TEC}{f_1^2} + n_1\lambda_1 \tag{2.20}$$

$$L_2 = P_0 - \frac{TEC}{f_2^2} + n_2\lambda_2 \tag{2.21}$$

where P_1 and P_2 are the pseudoranges from the precise P code, L_1 and L_2 represent carrier phases of the signals in meters. P_0 is for an ionosphere free pseudorange while n and λ represent the integer ambiguity and the carrier wavelength respectively. The term ε represents the dispersive component of the receiver and satellite hardware errors.

Taking the difference between the observables for both phase and code ranges, two equations for TEC can be obtained:

$$\frac{P_1 - P_2}{f_1^{-2} - f_2^{-2}} = TEC + \frac{\varepsilon_1 - \varepsilon_2}{f_1^{-2} - f_2^{-2}} \tag{2.22}$$

$$\frac{L_1 - L_2}{f_1^{-2} - f_2^{-2}} = -TEC + \frac{n_1\lambda_1 - n_2\lambda_2}{f_1^{-2} - f_2^{-2}} \tag{2.23}$$

Now TEC values can be obtained from Equation 2.22 along with a noise term and Equation 2.23 with an offset term from the integer ambiguity associated with it. The integer ambiguity is constant as long as the satellite is visible (except during cycle slips, events of large or sudden changes). The offsets associated to large cycle slips are easily detectable. TEC can be calculated by substituting Equation 2.23 into Equation 2.22. In addition, the offset between such slips is equivalent to a weighted mean of the differences between Equations 2.22 and 2.23 and the weights are correlated with the signal to noise ratio (Allain and Mitchell, 2009).

By obtaining the TEC measurements from the above equations, reconstruction of

the ionospheric electron density profiles and imaging of the ionospheric TEC map is possible through tomography.

2.2 Ionosonde

An ionosonde is a HF pulsed radar used to monitor the bottomside ionosphere. It sends a sweep of frequencies to the ionosphere and records the time delay for each frequency reflected back at different heights of the ionosphere. The frequencies can range from below 0.1 MHz to 30 MHz or more with a sweep duration from a few seconds to a few minutes (Davies, 1990). Ionosondes are classified according to two main modulation methods used: pulse and chirp. The pulse ionosondes transmit short pulses of approximately $100\mu s$ at a high peak power of 1 - 30 KW. The receiver outputs of the pulse ionosondes are peaks in the time domain. The chirp ionosondes transmit a long pulse of 0.5 - 1 seconds at a lower peak power (10 - 100 Watts). The receiver outputs are peaks in the frequency domain. The South African ionosonde network uses pulse ionosondes.

The output of an ionosonde record is displayed as an ionogram which is a plot of virtual height against operating frequency (McNamara, 1991, page 20). Virtual height is the distance that a radio wave, at a specified frequency, would have traveled in half the elapsed time, had it traveled at the speed of light in free space (McNamara, 1991, page 52). Assuming the ionosonde sounding waves travel at the speed of light c, the virtual height h' is related to the delay time as (Davies, 1990, page 95):

$$h' = \frac{c\tau}{2} \tag{2.24}$$

where τ is the time delay of the reflected wave and the factor 2 accounts for the fact that the pulses make a round trip. The frequencies measured at different heights of the ionospheric layers are related to the corresponding electron density as given in

Equation 1.2. During reflection the ionosphere's electron plasma frequency matches the vertically reflected O-mode radio frequencies. Figure 2.3 shows a display of a typical ionogram. The ionogram displays two traces of the radio waves, the ordinary

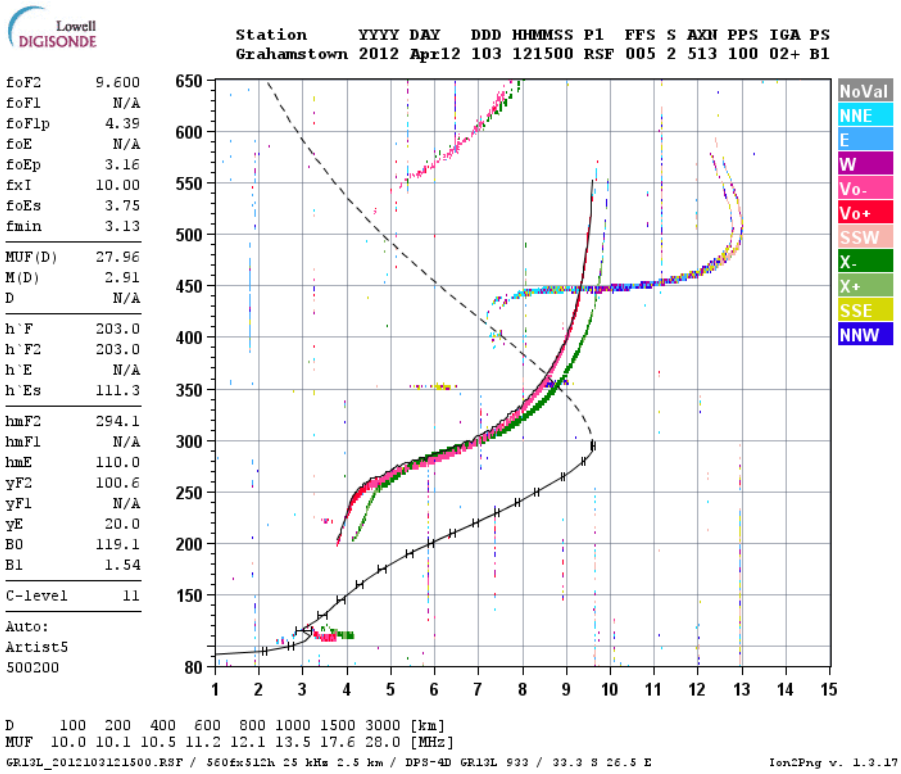

Figure 2.3: This figure illustrates how a typical ionogram displays the ionospheric parameters whose magnitudes are shown on the left outer side of the ionogram. This ionogram shows the peak frequencies and corresponding heights of the ionosphere on 12 April 2012 at the Grahamstown ionosonde station.

(O) and extraordinary (X) modes, which in Figure 2.3 are in pink and green colors respectively (Liu et al., 2004). The vertical asymptotes for the O-mode and X-mode waves traces in the ionogram represent the critical frequencies foF2 and fxF2 respectively (McNamara, 1991). The critical (penetration) frequency of the ionospheric F_2 layer (foF2) is the highest frequency on the O-mode trace.

The South African Ionosonde Network currently comprises of four digisonde portable

sounder (DPS) ionosondes located at Grahamstown (Eastern Cape, 33.3°S, 26.5°E), Louisvale (Northern Cape, 28.5°S, 21.2°E), Madimbo (Limpopo, 22.4°S, 30.9°E) and Hermanus (Western Cape, 34.4°S, 19.2°E). The sampling period of the ionosondes is 15 minutes. Figure 2.4 shows the geographic location of the ionosonde stations used in this study.

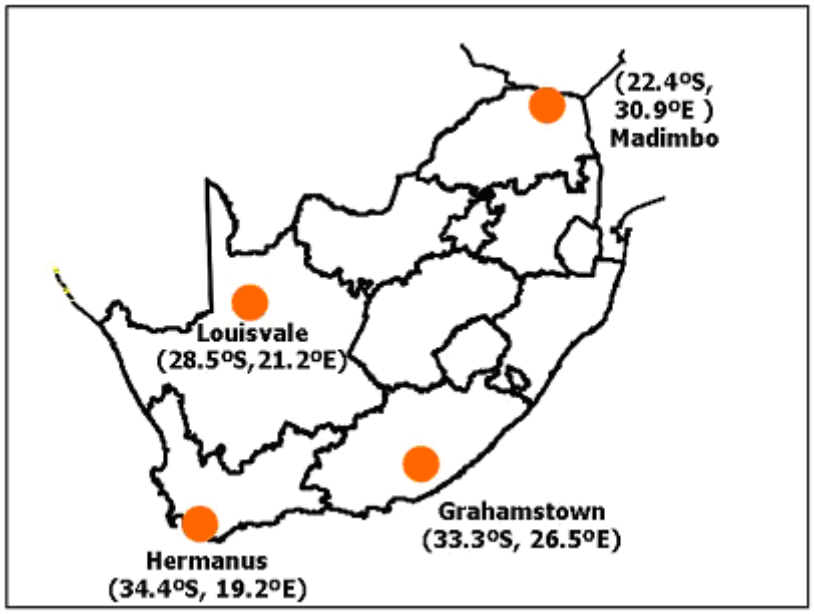

Figure 2.4: Locations of available South African ionosonde stations.

There are two main shortcomings with ionosonde measurements. One is their inability to cover the whole range of the ionosphere (i.e. they measure only the bottomside of the ionosphere). The second is, unlike GPS measurements, ionosonde measurements are very localised because they measure the ionosphere directly above their geographical location. Nevertheless, ionosondes are believed to provide true measurements of the bottomside ionosphere. Measurements of the ionosonde from the Digisondes deployed in South Africa, such as the critical frequencies, TEC and corresponding heights of reflection layers can be extracted using the Standard Archiving

Output (SAO)-Explorer software. SAO-Explorer is a data analysis tool which enables manual editing and viewing of autoscaled ionograms and derived ionospheric characteristics (Reinisch et al., 2004). The researches main focus was to see if the general trend of the diurnal NmF2 values extracted from the 1D electron density profiles derived from ionosonde measurements can be represented by the MIDAS reconstructed NmF2 values. For this reason, autoscaled profiles of the derived ionospheric characteristics from the four South African ionosonde stations were used for validation of the MIDAS (to be discussed in chapter 4) derived profiles.

Chapter 3

Ionospheric Tomography

3.1 Introduction to Ionospheric Tomography

In many applications it is worth examining the internal structure of an object or body without opening it (Fehmers, 1996). For example, physicians want to see the inside of human body, geologists need to measure the cross-section of the Earth or even need to explore natural resources buried deep down inside the Earth's surface. A non-invasive technique to study these cross-sections is grouped under a class of problem solving called tomography (Fehmers, 1996). Tomo comes from a Greek word $\tau o \mu \eta$, meaning slice and the term tomography represents the process by which a two-dimensional image of a section or a slice of a three-dimensional body is obtained (Rose, 2011; Fehmers, 1996; Bust and Mitchell, 2008). Fehmers (1996) explained that all tomographic techniques are based on waves, be it electromagnetic or acoustic in nature. As these waves traverse through an object they are emitted, absorbed, transmitted, reflected, refracted and diffracted, or combinations of these, by the internal structures of the object; therefore, tomography can be of emission, reflection, transmission and diffraction. In addition, better reconstruction of the object of study is attained if the number of viewing directions is large (Fehmers, 1996).

The tomography technique was first applied to radio astronomy in 1956 (Bracewell, 1956). This work explained the idea of creating a complete and full reconstruction of the true distribution of a celestial object from a finite number of scans in all possible directions. Austen et al. (1986) proposed the first application of the tomographic technique to ionospheric imaging. Then two years later Austen et al. (1988) suggested radio tomography methods based on measurements of satellite signals. This work demonstrated the feasibility of reconstructing ionospheric electron density from total electron content (TEC) measurements measured using LEO satellites and ground receivers. It also pointed out that the reconstruction results become reasonable, if as many rays cross the given region as possible over a wide range of angles. The application of tomography to the ionosphere where dual frequency trans-ionospheric radio waves are analyzed and inverted to produce two-dimensional images of electron concentration, falls into the category of ray tomography (Bust and Mitchell, 2006). A review of tomography algorithms based on TEC measurements is presented in Kunitsyn and Tereshchenko (2003), pages 21-32.

3.2 Ionospheric Tomography

Ionospheric tomography is the reconstruction of vertical electron density profiles of the ionosphere from a network of line integrals. The phase changes of GPS signals crossing the ionosphere are recorded either by ground- or space-based receivers. In the case of conventional tomography the measurements of the line integrals are made from dual-frequency trans-ionospheric signals along many intersecting ray paths between Low Earth Orbiting (LEO) satellites and a chain of ground-based receivers (Dear and Mitchell, 2007). A study by Yeh and Raymund (1991), showed that the geometry of LEO satellite-to-Earth measurements only allows observations over a

limited number of viewing angles that would impose limitations to ionospheric tomography. The authors further explained that orientations of satellite-to-receiver ray paths are biased in a vertical direction with no ray paths running horizontally through the ionosphere owing to the curvature of the Earth, which results in a poorly defined vertical electron concentration gradient taken by TEC measurements alone. This indicated that the reconstruction of the tomographic images using conventional inversion algorithms alone is not straightforward and difficult to implement; therefore, tomographic problems need to consider the geometrical constraints in developing inversion algorithms. Howe et al. (1998) reviewed the tomographic algorithms that use polar-orbiting satellites as signal transmitters and a chain of ground receivers to reconstruct two-dimensional maps of the electron density. It pointed out that tomographic solutions using data from a network of LEO satellites and ground receivers alone was not enough to study the ionosphere; therefore, the problem demanded better solutions. There were several algorithm solutions developed to overcome this problem, but the algorithm by Fremouw et al. (1992) was widely used.

The inversion algorithm in Fremouw et al. (1992), which was adapted from geophysical applications, uses a set of vertical orthonormal functions created from ionospheric models to image the vertical profile and a power law spectrum to select the horizontal structures from a Fourier basis. The algorithm in Mitchell and Spencer (2003), a three-dimensional time varying ionospheric tomographic imaging, is an extension of a two-dimensional static inversion algorithm developed by Fremouw et al. (1992). The third version of this algorithm, which uses a voxel based three-dimensional spatial grid instead of spherical harmonics for the horizontal structures, is used in this study.

After GPS satellites were made available, GPS signals recorded at ground- or space-(onboard LEO satellites) based receivers became an important tool for ionospheric measurements. GPS signals can also be recorded by receivers onboard LEO satellites in a technique called radio-occultation. Radio-occultation is a technique where receivers onboard LEO satellites monitor the phase changes of GPS signals. The work of Hajj et al. (2004), suggested using GPS to LEO satellite measurements in a tomographic problem to provide the missing horizontal rays and improve the vertical resolution. Hajj et al. (2004) also added that LEO satellites provide measurements over oceans and into remote polar caps; therefore, enabling the study of the ionosphere on a global-scale. The works of other authors using GPS receivers on board satellites and the LEO missions used was reviewed by Bust and Mitchell (2006).

After tomographic imaging had proved to be a successful technique in the scientific study, it became a viable tool for practical applications (Bust and Mitchell, 2006). Applications of tomography in other fields include computerized tomography in medicine in which two-dimensional images of a human body can be reconstructed from measurements of the attenuation of X-rays crossing through it from many different angles (Cilliers et al., 2004). Another is seismic tomography, used for imaging the geophysical structure of the Earth's subsurface. The methods and equipment for seismic tomography are well developed. In seismic tomography measurements of propagation times of seismic waves are used for reconstructing seismic slowness, a quantity reciprocal of the wave velocity (Kunitsyn and Tereshchenko, 2003, page 21).

TEC measurements are basic parameters for ionospheric tomography; however, individual TEC measurements provide no information about the spatial variation of the electron concentration along the ray path. Yet, tomographic algorithms can ingest these linear integral data to produce two- or three- dimensional images of the

electron density of the ionosphere. In addition, tomographic systems are advantageous over conventional in situ point measurements in that many independent data can be obtained with relatively small instruments. This means the amount of data can grow approximately quadratically with an increase in instruments (Howe et al., 1998).

The quality of ionospheric tomographic inversions depends mainly on the density of the ray paths which in turn depends on the number of intersecting ray paths between transmitters and receivers over a wide range of angles (Kunitsyn and Tereshchenko, 2003). Studies showed that data collected from an uneven distribution of ground receivers impacts the quality of the reconstructed images as the case when sparse data is used. According to the studies of Zapfe et al. (2006) and Materassi and Mitchell (2005), an uneven distribution of receivers lead to low information content of the dataset. The reconstruction based on such an input dataset severely impacts the resolution images of electron density and TEC. Zapfe et al. (2006) indicated that dense and even receiver distribution were found to give low maximum absolute percentage errors while uneven receiver distribution was associated with maximum absolute percentage errors.

The development of global, three-dimensional, time-varying algorithms made use of the GPS constellation of satellites as data sources. Imaging in three dimensions is basically the same as in two dimensions, except there are some complications due to the sparseness of data coverage that have to be overcome. Due to the fact that iterative pixel-based methods such as ART and MART (discussed in Bust and Mitchell (2008)) do not couple the changes introduced to one voxel across to other voxels that may not have intersections with measurements, make them less important in solving the problem (Bust and Mitchell, 2008). Therefore, defining a three-dimensional grid of voxels, each bounded in latitude, longitude and altitude is required for the

inversion process. The contribution of the individual straight-line elements of the signal path between satellites and receivers at each intersected voxel is computed from the knowledge of the satellites and receivers locations (Bust and Mitchell, 2008; Mitchell and Spencer, 2003). Then the unknown electron concentration is required to be constant within each voxel and contained in the column vector x_i. The nature of the problem can be written as:

$$Ax_i = b \tag{3.1}$$

where the matrix A transforms the electron density to the form and location of the observations and b are the set of observed TEC values (Bust and Mitchell, 2008). This approach of solving the tomographic problem is used by MIDAS and is discussed in the next chapter.

Another approach to three dimensional tomographic imaging of the ionosphere is data assimilation. Data assimilation is basically a model specification and a prediction technique which uses data to improve the exactness of the model (Bust and Mitchell, 2008). This technique is based on first-principle physics models which comprises of multiple species of ions and neutrals all defined by coupled Maxwell equations of the individual continuity, momentum, and energy equations (Bust and Mitchell, 2008). Examples of the algorithms that fall in this group are IDA3D (Ionospheric Data Assimilation 3-Dimensional), developed by the Applied Research Laboratories, University of Texas at Austin; GAIM (Global Assimilation of Ionospheric Measurements), developed by Utah State University (USU) and another version of GAIM developed by the University of Southern California and Jet Propulsion Laboratory (USC/JPL). A review of these imaging algorithms is given by Bust and Mitchell (2008), Opperman (2008) and Bust and Mitchell (2006).

3.3 Ionospheric Models

The paucity of input data is a major problem that all tomographic inversion techniques experience. The best solution to alleviate this problem so that the inversion solutions are reasonable, is to incorporate *a priori* information about ionospheric parameters of interest. One application of empirical models (such as International Reference Ionosphere (IRI) model) and physics based functions (e.g Chapman function) is to provide background information of the ionosphere in regions where the primary source of data is absent. A brief discussion of ionospheric models such as the IRI model and Chapman profile is presented here. MIDAS algorithm uses either Chapman functions or the IRI model to create orthogonal empirical functions (EOFs) from a set of radial profiles of the ionospheric electron density distribution (Spencer, 2008).

3.3.1 Chapman Profile

For an isothermal atmospheric model the rate of photoionization decreases exponentially with increasing altitude. The general equation of Chapman function for electron density profile is:

$$N = N_o e^{k\left[1-z-sec\chi e^{-z}\right]} \tag{3.2}$$

where χ is the solar zenith angle; N is electron density; k is a constant with values $k = 1$ for Chapman beta layer and $k = \frac{1}{2}$ for Chapman alpha layer and z is the reduced height given by

$$z = \frac{h - h_m}{H} \tag{3.3}$$

in which h and h_m are the height and peak density height respectively, and H is the scale height. The two Chapman layers are associated to the heights where the dominant ion loss reactions occur; i.e the square law loss formula (Chapman α,

dominant in E region) and linear loss formula (Chapman β, dominant in the F_2 region). A detailed discussion of this topic can be found in Rishbeth and Garriott (1969), pages 89-94. The Chapman functions predict maximum electron densities of the E- and F-layers from which the empirical orthogonal functions are derived. MIDAS uses the Chapman beta function to create EOFs.

3.3.2 The International Reference Ionosphere (IRI) Model

The international reference ionosphere (IRI) is a standard for the specification of plasma parameters in the Earth's ionosphere (Bilitza, 2001). The empirical model is continuously updated by a joint working group of the Committee on Space Research (COSPAR) and the International Union of Radio Science (URSI) since its first establishment in 1969 leading to a number of versions.

The IRI model is an empirical model because it uses most of the available and reliable data sources. The advantage of IRI, being an empirical model, is that the changing theoretical understanding of the ionospheric plasma processes does not affect the model results (an example of this scenario is discussed in Bilitza (2001)). However, the model's strong dependency on the underlying data base makes it unreliable in areas of less or no contribution of input to the data base. For this reason, the results become less reliable in the Southern hemisphere due to the scarcity or more often unavailability of ionosonde networks.

MIDAS uses IRI95 model coefficients to create EOFs and the EOFs are used to constrain the MIDAS reconstructed electron density, in the vertical direction. In addition, there is a provision within MIDAS to use electron densities from the IRI95 model as *a priori* information to start the inversion process.

3.4 Summary

In general, tomographic techniques reconstruct electron density profiles from the observable GPS parameter TEC. In places where there is scarcity of data, the only way to fill the gap is to use ionospheric models such as Chapman and IRI so that a reasonable solution can be found.

In the next chapter, an inversion algorithm called MIDAS which reconstructs time varying three-dimensional electron density profiles from GPS TEC measurements is presented. As ionosonde electron density profiles are one-dimensional, a one-dimensional electron density profile produced from MIDAS is compared with a similar profile from ionosondes at the selected ionosonde stations.

Chapter 4

MIDAS

The Multi-Instrument Data Analysis System (MIDAS) is a four-dimensional time-dependent tomographic algorithm (Muella et al., 2011). MIDAS, developed by an Invert group at the University of Bath in the United Kingdom, was adopted from the work of Fremouw et al. (1992). This algorithm, like other inversion techniques, ingests data from instruments such as GPS and ionosondes. A number of ionospheric studies have been made using MIDAS; for example, Muella et al. (2011) used MIDAS to image the equatorial and low latitude ionosphere over Brazil, Materassi and Mitchell (2005) used it to image the equatorial anomaly and Katamzi (2008) and Cilliers et al. (2004) used it to characterize the ionosphere over the South African region. The work of Katamzi (2008) used MIDAS1 (the first version) to reconstruct one-dimensional electron density profiles using data from the then available South African network of ground-based GPS receivers and verify the profiles with an ionosonde-derived electron density profile from the Grahamstown, South African station. This research is different from the Katamzi (2008) work in the following aspects: MIDAS3, which is used in this research, uses a voxel based grid of the ionosphere to constrain the horizontal variation of the electron density profiles whereas MIDAS1 uses spherical harmonics to modulate the variation in the horizontal structures. MIDAS3 also assumes a linear variation of the electron

density profiles through out the grid within the time window of inversion. In addition, MIDAS3 uses two inversion methods which are discussed in section 4.2. The other important difference with this research is that roughly 53 GPS receivers (as opposed to the 10 receivers used in the Katamzi (2008) work) were used since the GPS network of receivers has expanded significantly over the past few years. Due to the four dimensional nature of the GPS observations in space and time and the fact that the ionosphere cannot be assumed to be constant in the time interval required to obtain sufficient measurements, developing a four dimensional approach to the inverse problem was crucial. For this reason MIDAS was upgraded to a four-dimensional version to address this problem (Spencer, 2008). A brief description of the MIDAS algorithm is presented in the following section and more details can be obtained from a technical report by Spencer (2008), and publications by Spencer and Mitchell (2007), Materassi and Mitchell (2005) and Mitchell and Spencer (2003).

4.1 Basics of MIDAS Algorithm

The first stage of the inversion is to divide the ionospheric region of interest into a grid of three-dimensional volume pixels, called voxels, bounded in latitude, longitude and altitude. Latitudes and longitudes measured in degrees respectively and altitude measured in kilometers above the surface of the Earth. MIDAS uses a function based on a MATLAB function called *ndgrid* to enable an efficient construction of grids of the three-dimensional region of interest. The schematic representation of the gridded ionosphere is shown in Figure 4.1.

In many inverse problems, including the classical tomographic problem, data are measured as line integrals along rays (Spencer, 2008); hence, the solution for such

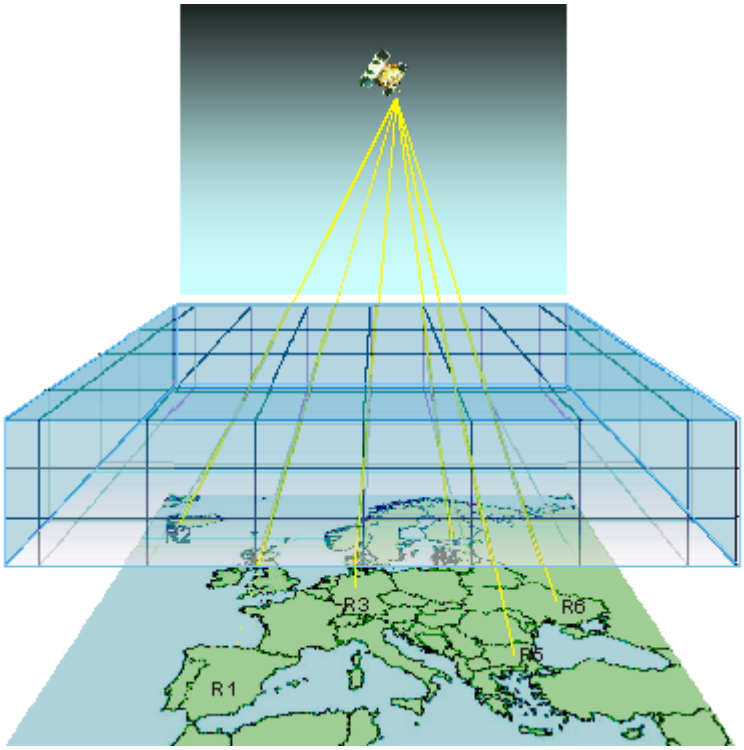

Figure 4.1: Figure shows how the ionosphere is gridded into three-dimensional boxes called voxels and the ray paths that cross this region from a single satellite to few ground receivers. *after* Mitchell (2005)

problem requires that line integrals be represented in matrix form:

$$Hx = z \tag{4.1}$$

where, H defines a decomposed matrix of the per voxel path integral contribution for each ray at each vertex, x is the unknown electron densities and, z are the measurements of the line integrals obtained from receivers. Now the problem becomes one of inverting H, a highly singular matrix which depends on the geometry of the ionospheric region of interest setup for inversion, to solve for the unknowns, x. MIDAS incorporates a tool to compute the matrix H which takes different shapes depending on the nature of the problem. The signal rays can propagate along straight lines in which case the problem becomes linear or can refract while passing through the ionosphere whereby the problem becomes non-linear depending on the dependence

of the matrix H on the field x at each vertex (Spencer, 2008). Based on the nature of the problem, a linear approach for a linear problem and a non-linear or iterative approach for a non-linear problem must be used.

The ray tracing mechanism in inverse problems allows three possible relationships between the field, x, and the refractive index, μ, of the medium upon which the propagation path of the rays depend. These are:

$$\mu^2 = 1 - Ax \tag{4.2}$$

$$\mu = \frac{A}{x} \tag{4.3}$$

$$\mu = 1 \tag{4.4}$$

Equations 4.2, 4.3, and 4.4 represent radio wave propagation with no magnetic field, acoustic propagation and a no refraction (straight line) ray respectively. A is a user defined constant. Analytical solution of the path of a ray through a linear gradient medium is possible for these three relationships. The voxels contained in a three dimensional volume are further decomposed into 6 tetrahedra by a ray tracer which propagates a ray through them. The contribution to the path integral at each vertex of tetrahedron from the ray paths passing through the grid is summed and returned as the H matrix for inversion. A detailed discussion on a selection criterion for rays, contained in each voxel, traversing the region of interest between transmitters and receivers to be considered as valid or otherwise neglected, is presented in Spencer (2008). Figure 4.2 shows the decomposition of a cube into tetrahedrons.

The output matrix of the ray tracer, H, contains a per vertex decomposition of the ray path where each row corresponds to a ray and each column to a vertex. The ray path matrix is constructed in such a way that the sum of H along all the columns is the ray path integral of the ray. In the case of linear propagation where refractive index of the medium is unity, the optical path of the ray is simply the geometrical

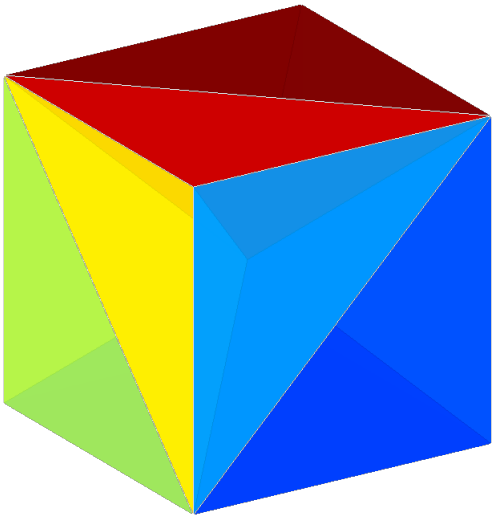

Figure 4.2: Tetrahedral decomposition of a cube by the ray tracer. *after* Spencer (2008)

distance.

Solving the inversion problem in Equation 4.1 as it stands is difficult as the matrix H is highly singular. A number of tomographic inversion algorithms developed to solve the inversion problem follow different methods, but they all share a common procedure, converting the rectangular matrix to an invertible matrix. Transformation of a problem from one basis set to another basis set is an important technique in inverting poorly constrained inverse problems. This is done using mapping functions. Mapping is accomplished by constructing a matrix, M, that transforms the inverse problem such as Equation 4.1 defined in one basis set to:

$$(HM)X = z \tag{4.5}$$

in which X, represents the solution in the transformed basis set. These mapping functions are often large matrices constructed in different dimensions which takes account of the measurement geometry or the nature of the medium being reconstructed. MIDAS uses a number of functions to construct mapping matrices; the

International Reference Ionosphere (IRI) model, Chapman functions and a convolution map, to mention a few. Chapman functions and the IRI model are used to create empirical orthogonal functions (EOFs) from a set of radial profiles of the ionospheric electron density distribution. To carry out mathematical operations of the mapping matrices as they are is simply not possible since the matrices involved would be too large to store; therefore, it is beneficial to store maps in their decomposed cell array notation. MIDAS uses the convolution map to create a cell decomposed array of the mapping matrices for the three dimensional grid given a cell decomposed kernel. For more information about convolution map, a MATLAB based mathematical transform, the reader is referred to the Mathworks website (http://www.mathworks.com/help/signal/ref/convmtx.html).

The MIDAS inversion technique is based on the use of orthogonal functions, an idea substantiated in Fremouw et al. (1992). Since the inversion problem is underdetermined a model or mathematical based orthonormal functions is/are used to compensate for the missing information to allow for the reconstruction of a spatially continuous electron density field (Bust and Mitchell, 2008). According to Materassi and Mitchell (2005), MIDAS inversion requires functions to define the vertical profile of electron concentration. This kind of inversion problem is highly underdetermined in general (Fehmers, 1996) and the requirement for selecting the profiles is to request prior hypothesis (Materassi and Mitchell, 2005). The authors also added that the free electron density, in MIDAS inversion, is assumed to be a linear combination of a given set of three-dimensional EOFs.

Along with the three-dimensional volume grid a time mapping is also added in constructing the mapping function. The mapping matrix uses a set of EOFs to map the radial profile and a voxel based grid bounded in latitude and longitude horizontally and a time axis of reconstruction which results in time-windows of the

inversion.

All phase observations, z, within the constructed time varying 3D volume are incorporated into an observation matrix, H, so that they can be mapped using the mapping matrix M defined above. An efficient usage of all the available information in the data can be made by differencing the rows of H and z for all pairs of rays obtained from each satellite/receiver for which signal lock is maintained. This problem is then converted into a set of normal equations, which involves a regularization matrix R to force quadratic smoothing on the solution both in horizontal space and in time, as:

$$H^T H + cR^T R = H^T z \tag{4.6}$$

where c is a user defined regularization constant.

There are a number of sparse inverse methods included in MATLAB to solve the normalized tomographic problems such as Equation 4.6. In MIDAS the user can choose to use either *minres* (Minimum Residual Method) or *quadprog* (Quadratic Programming Method), both MATLAB based methods, to do the inversion.

4.2 MIDAS Inversion Methods

4.2.1 Minimum Residual Method

Minimum residual method attempts to find a minimum norm residual solution x to a system of linear equations such as:

$$Ax = b \tag{4.7}$$

As a requirement the coefficient matrix A needs to be a large, square and sparse matrix and the column vector b must have a length equal to the row or column length

of matrix A. In MIDAS, the *minres* method solution to Equation 4.7 in the transformed basis can have negative values which leads to negative electron density profiles. The *minres* method does not include an option whereby one can constraint the inversion process to get a non-negative solution. For this reason *quadprog* method was used to solve the inversion problem in this study. For more information about the *minres* method refer to the Mathworks website (http://www.mathworks.com/help/matlab/ref/minres.html).

4.2.2 Quadratic Programming

Quadratic programming method solves quadratic problems of the form:

$$\begin{aligned} &\min(\frac{1}{2}X^T HX + f^T X) \\ &\text{subject to :} \quad HX \leq b \end{aligned} \tag{4.8}$$

in such a way that the method returns a vector X which minimizes the above problem. The method requires H to be a positive definite in order for the problem to have a finite minimum. This method takes in several inputs such as upper and lower bounds, equality and inequality constraints to impose a constraint on the solution to the inversion problem. In order for the solution to the inverse problem in the transformed basis, X, to be non-negative, the *quadprog* method imposes a constraint to the inversion problem via the inequality constraint:

$$MX \geq 0 \tag{4.9}$$

which guarantees a positive set of vectors of EOF coefficients. Then, the final solution in the original basis, i.e the electron density profiles can be obtained by:

$$x = MX \tag{4.10}$$

4.3 MIDAS Inversion Procedures

The GPS data used in this study were obtained from a network of permanent continuously operating Global Navigation Satellite System (GNSS) base stations distributed throughout South Africa at approximately 100 km to 300 km spacing. This network, called Trignet, is installed, maintained and operated by Chief Directorate Surveys and Mapping (http://www.trignet.co.za/). The GPS data collected at 30 second intervals were used as an input to MIDAS to make the inversion. MIDAS uses observation data from a time-window around the inversion, so in this study a time-window of three hours and forty five minutes with a time step of fifteen minutes is used. A time varying three dimensional spatial grid covering the South African region was used in this inversion. The horizontal boundaries were latitude ranging from 15° S to 39° S and longitude ranging from 11.5° E to 40° E, with a spatial resolution of 2° each. Vertically, an altitude ranged from 25 km to 1200 km for which the lower region, 25 km to 500 km, was in steps of 25 km and the upper region, above 500 km, was in steps of 100 km.

The choice of EOFs in MIDAS determines the results that one can obtain from the time-dependent inversion process. The use of EOFs is vital because they reduce the dimensionality of a dataset consisting of multiple intercorrelated variables, while preserving the majority of the variation influenced by independent processes to the maximum level. This is possible by the decomposition of the original dataset into a set of uncorrelated and ordered base functions and their associated coefficients using orthogonal transformation (Ercha et al., 2012). The authors further explain that the orthogonal base functions are not predetermined artificially, but derived from the original dataset via the process of decomposition. As a result, each succeeding base function contains the variance in the original variables as much as possible which makes the EOF series converge rapidly; therefore, few EOF base functions

can describe the internal characteristic of the data while the majority of the variance in the original data can be best represented. This property of EOFs makes it a useful tool in inversion algorithms whereby they impose a restriction in the electron density profiles in the vertical sense.

In MIDAS the electron densities are generally assumed to change linearly in each voxel over the time-window of inversion. The right tuning of the number of EOFs bring about improved optimization of electron density profiles. In this study Chapman functions and IRI95 model were used to form the empirical orthogonal functions. By varying the number of EOFs better results can be obtained from the inversion process. What follows is a summary of the procedures that were followed during the inversion to obtain the results that are discussed in Chapter 5.

For EOFs constructed from the Chapman β function, the following parameters were used: peak heights ranging from 260 km to 320 km during day time and up to 360 km during night time and scale heights ranging from 60 km to 90 km. In addition, a grid of the ionosphere of interest bounded in latitude and longitude is inputted to get the three dimensional spatial set of EOFs which imposes the electron densities to be reconstructed from the inversion algorithm.

For EOFs constructed from the IRI95 model, the following parameters were used: number of required EOFs, date, and a specified spatial grid of the ionosphere as mentioned above. The EOFs were computed using singular value decomposition. The unknown electron densities are represented by these orthogonal vertical profiles, modulated horizontally via the voxel based grid of the ionosphere of interest bounded in latitude and longitude. The inversion was made for a different number of EOFs (i.e 1, 2, 3, 4 and more) and the results are discussed in Chapter 5.

After computing the EOFs using either the Chapman function or the IRI95 model,

the inversion process can be done with or without *a priori* information of the selected ionosphere. In this study, the starting estimate of the solution to the inversion was formed without *a priori* information of the ionosphere.

The profiles reconstructed from MIDAS using the two methods and the procedures mentioned above are presented in the following chapter. The 2-dimensional electron density contour plots were produced for randomly selected days in 2012. MIDAS reconstructed 1D electron density profiles were compared with ionosonde derived 1D electron density profiles. The verification was made only in terms of NmF2 values extracted from the 1D electron densities obtained from MIDAS reconstruction and ionosonde measurements.

Chapter 5

Results and Validation

Eight days in 2012 were randomly selected from each season for the validation of the MIDAS produced electron density profiles using ionosonde profiles. These 8 days were: 12 April and 12 May (representing autumn), 12 July and 12 August (representing winter), 12 October and 12 November (representing spring) and 12 and 21 December (representing summer). Dst indices for 6 days around each validation day were included in Figure 5.2 and the lowest Dst index from the 8 validation days was -32 nT. Since according to Loewe and Prolss (1997) Dst indices greater than -30 nT describe quiet geomagnetic conditions the 8 days were on quiet geomagnetic conditions.

Figure 5.1 and Table 5.1 illustrate locations and geographic coordinates of the approximately 53 GPS receivers used in this study. Since the spatial resolution of the inversion grid is 2° by 2°, MIDAS profiles from the closest grid point to an ionosonde location were used for validation. Table 5.2 presents the coordinates of the grid points closest to each South African ionosonde station as well as the distances (in km) between the grid points and the ionosonde stations, refer to Figure 2.4 for ionosonde stations coordinates.

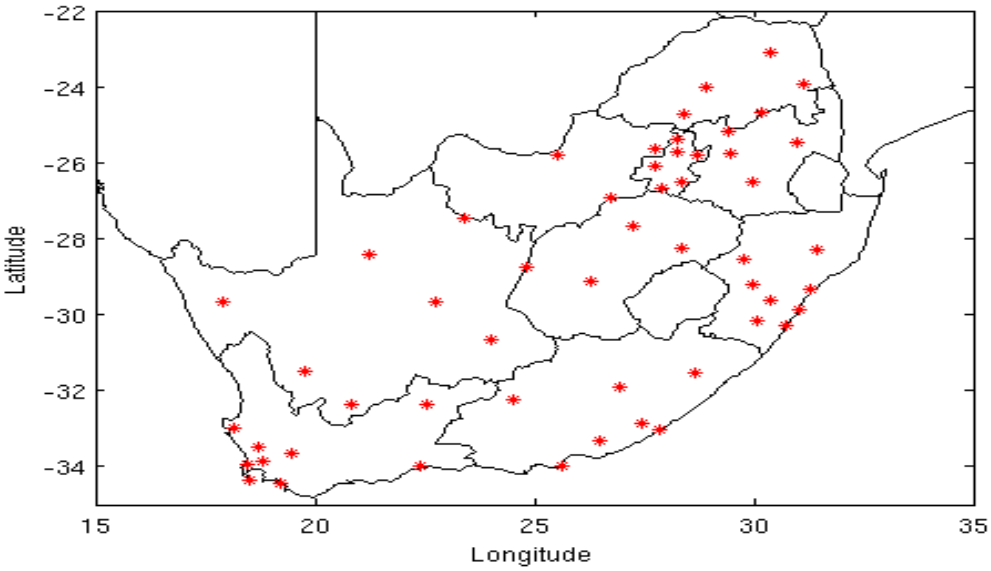

Figure 5.1: Geographic locations of the 53 GPS receiver stations used in this study.

Table 5.1: Geographical locations of GPS stations.

Station Name	**Station Code**	**Latitude °S**	**Longitude °E**
Bethelehem	BETH	28.25	28.33
Bloemfontein	BFTN	29.1	26.29
Bisho	BISO	32.86	27.42
Brits	BRIT	25.63	27.78
Bronkhorstspruit	BRNK	25.80	28.73
Beaufortwest	BWES	32.34	22.57
Calvinia	CALV	31.48	19.76
Capepoint	CPNT	34.35	18.48
CapeTown	CTWN	33.95	18.46
Deaar	DEAR	30.66	23.99
Durban	DRBA	29.85	31.02

Continued on next page

Table 5.1 – *Continued from previous table*

Station Name	**Station Code**	**Latitude °S**	**Longitude °E**
Eastlondon	ELDN	33.03	27.82
Ermelo	EMLO	26.49	29.98
Groblersdal	GDAL	25.16	29.41
George	GEOR	34.00	22.38
Grahamstown	GRHM	33.32	26.51
Graaff-Reinet	GRNT	32.24	24.53
Heidelberg	HEID	26.50	28.37
Hermanus	HNUS	34.42	19.22
Ixopo	IXOP	30.14	30.06
Kimberley	KLEY	28.73	24.80
Kuruman	KMAN	27.46	23.43
Krugerdrop	KRUG	26.08	27.76
Kroonstad	KSTD	27.66	27.24
Langebaan	LGBN	32.97	18.15
Ladysmith	LSMH	28.55	29.78
Malmesbury	MALM	33.46	18.73
Middleburg	MBRG	25.77	29.45
Mafikeng	MFKG	25.80	25.54
Mooiriver	MRIV	29.21	29.98
Nelspruit	NSPT	25.47	30.97
Nylstroom	NYLS	24.70	27.40
Orkney	OKNY	26.93	26.73
Phalaborwa	PBWA	23.95	31.13
Port Elizabeth	PELB	33.98	25.61
Pietermburg	PMNG	29.60	30.38
Potgtrust	POTG	24.01	28.93
Pretoria	PRET	25.73	28.28

Continued on next page

Table 5.1 – *Continued from previous table*

Station Name	**Station Code**	**Latitude °S**	**Longitude °E**
Prieska	PSKA	29.66	22.74
Queenstown	QTWN	31.90	26.92
Springbok	SBOK	29.66	17.87
Scottsburgh	SCOT	30.28	30.75
Steelpoort	SPRT	24.67	30.18
Stellenbosch	STBS	33.84	18.83
Stanger	STNG	29.34	31.29
Sutherland	SUT1	32.38	20.81
Thohoyandou	TDOU	23.07	30.38
Temba	TEMP	25.38	28.27
Ulundi	ULDI	28.29	31.42
Umtata	UMTA	31.54	28.67
Upington	UPTA	28.40	21.25
Vereeniging	VERG	26.66	27.90
Worcester	WORC	33.64	19.44

Table 5.2: Distances between ionosonde locations and inversion grid points.

Location Coordinates	**Grahamstown**		**Hermanus**		**Louisvale**		**Madimbo**	
	Lat	Lon	Lat	Lon	Lat	Lon	Lat	Lon
Ionosonde coord.(°)	33.3	26.5	34.2	19.2	28.5	21.2	22.4	30.9
Grid point coord.(°)	33	25.55	35	19.5	29	21.5	23	31.5
Distance in (km)	115.9		94.9		64.7		94.3	

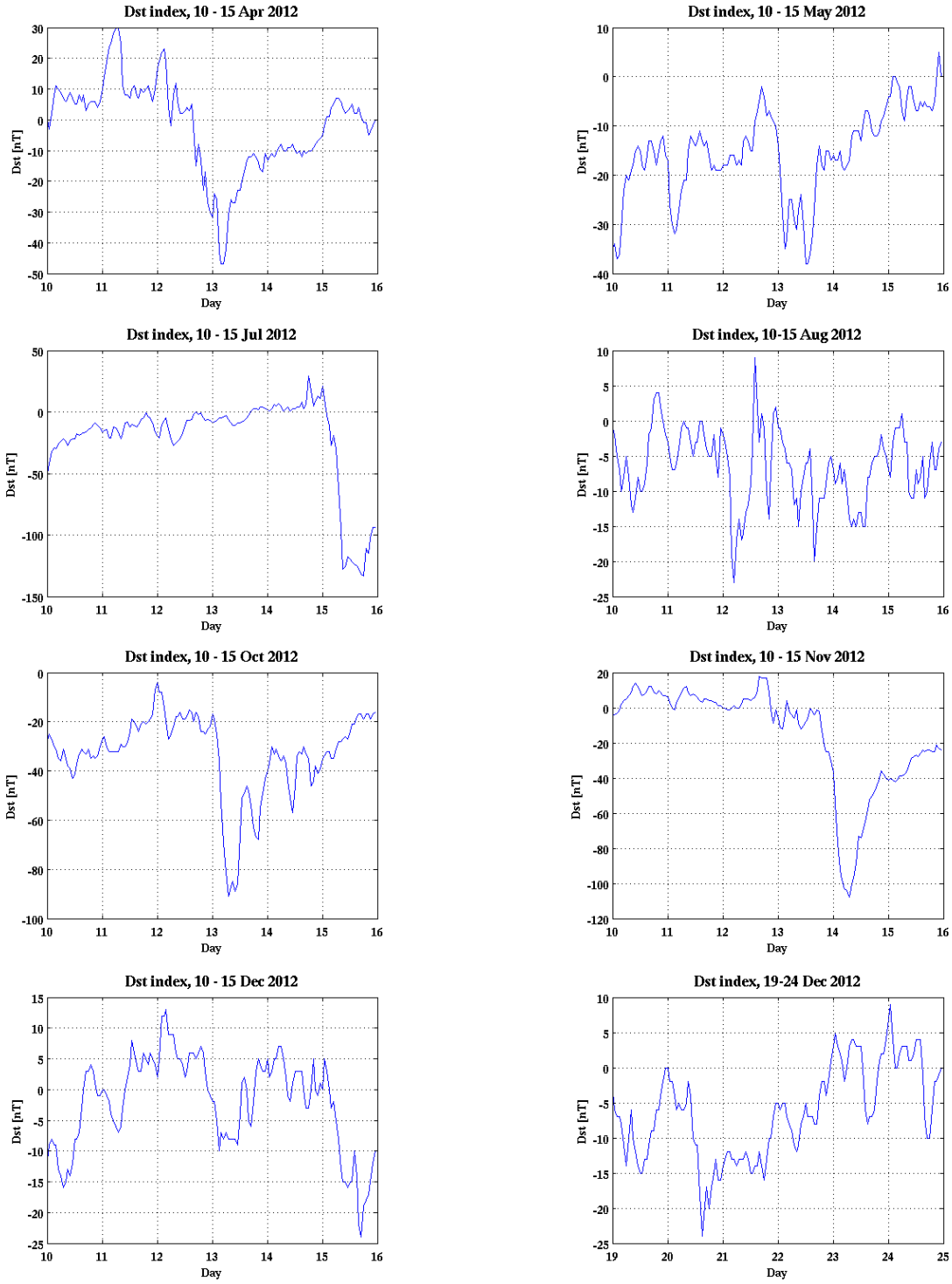

Figure 5.2: Dst indices on 6 days around each validation day.

5.1 Selection of EOFs

The MIDAS inversion process using the *minres* method produced negative electron density profiles. Figure 5.3 represents, as an example of results obtained from using the *minres* methods, MIDAS reconstructed 1D electron density profiles on 12 April 2012. Since clearly this is not acceptable, the *quadprog* method was used as an alternative method to produce the results used for validation.

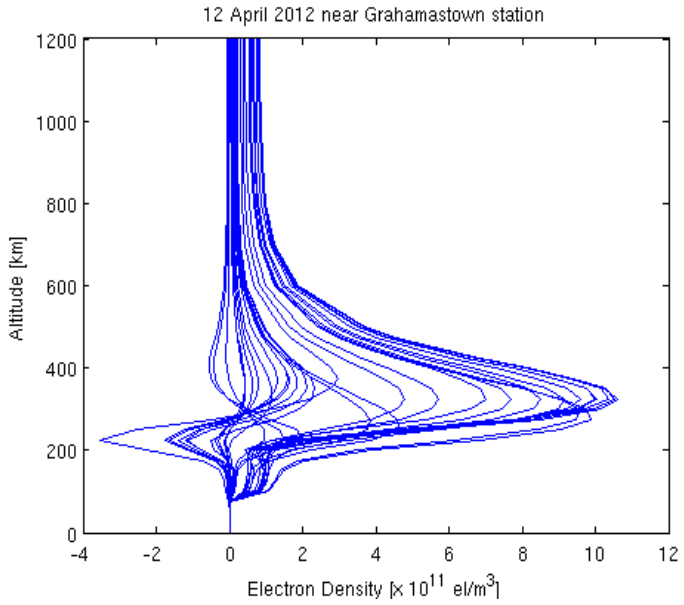

Figure 5.3: MIDAS reconstructed 1D electron density profiles on 12 April 2012 using the *minres* method. Each line represents a profile made in 15 minute interval of the day.

In order to decide whether the Chapman EOFs or the IRI EOFs produced the best results, a comparison of electron density profiles for each method was made. Figure 5.4 presents NmF2 diurnal trends from MIDAS reconstructed profiles using Chapman and IRI EOFs and ionosonde profiles for the (a) 12 April, (b) 12 May and (c) 12 July. From this figure, it is observed that the MIDAS reconstructed

NmF2 values using EOF1, ChapEOF1 and ChapEOF2 contained zero NmF2 values at the early and late evening hours of the days of the inversion. Chapman EOFs (ChapEOF1 and ChapEOF2) and IRI EOF1 methods produced zero NmF2 values for 7 out of 8 days near the Hermanus station; therefore, EOFs from the IRI were used to reconstruct electron density profiles from which the validation tests will be performed.

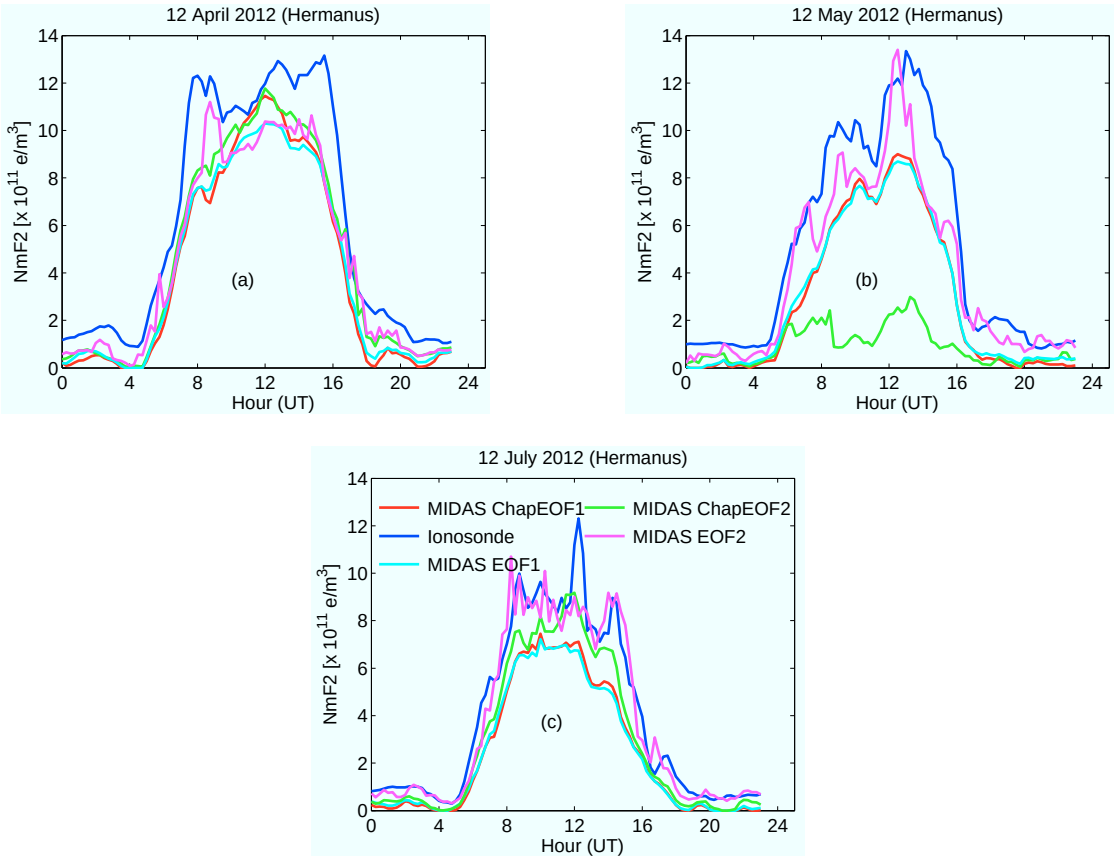

Figure 5.4: Comparisons of daily NmF2 values from the Hermanus ionosonde (blue) with MIDAS NmF2 values produced using Chapman EOFs (red and green) and IRI EOFs (cyan and magenta) on (a) 12 April, (b) 12 May and (c) 12 July 2012.

To decide on the number of IRI based EOFs to use for the reconstruction, a comparison of electron density profiles obtained from using 1 to 5 IRI EOFs was done. Figure 5.5 illustrates the electron density profile from the Grahamstown ionosonde used to compare with electron density profiles from a different number of IRI EOFs on 12 December 2012 at 10h00 UT. The results demonstrated that the MIDAS

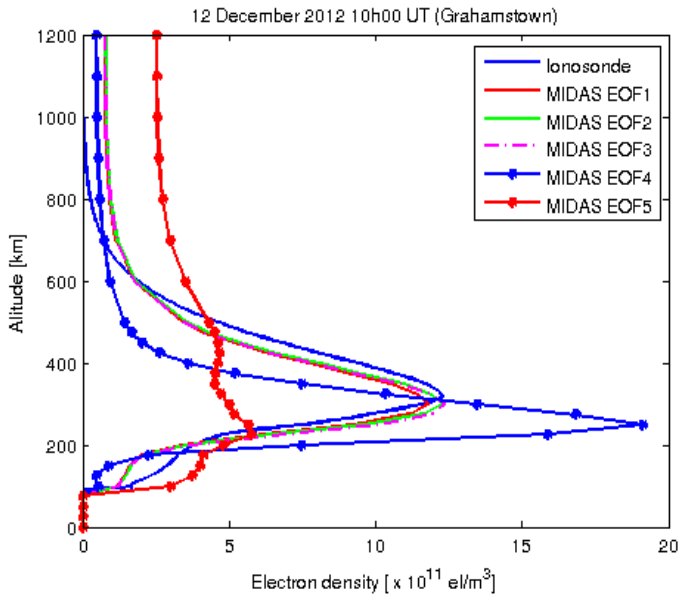

Figure 5.5: MIDAS reconstructed 1D electron density profiles using 1 IRI EOF (red-), 2 IRI EOFs (green -), 3 IRI EOFs (magneta -.-), 4 IRI EOFs (blue -*-) and 5 IRI EOFs (red -*-) compared to Grahamstown ionosonde profile (blue -) for 12 December 2012 at 10h00 UT.

reconstructed 1D electron density profiles using 4 EOFs highly overestimated the profile. In addition, the MIDAS reconstructed electron density profile using EOF5 not only underestimated the profile but also produced a double peak in the F_2 layer. It was concluded that the EOF2 produced the best results, therefore results from this method will be used for the validation process.

5.2 Results: 2D Electron Density Maps

Figure 5.6 illustrates the diurnal variation of the electron concentration on 12 April 2012. On this particular day the electron concentration peaked at 09h00 UT (pre-noon peak) and reduced substantially at local noon (10h00 UT) and stayed reduced for about an hour. The electron concentration then started to increase gradually until a post-noon peak occurred at around 14h00 UT. This phenomenon is also evident in the MIDAS and ionosonde NmF2 values at Madimbo ionosonde station (see Figure 5.13 (c)). This occurrence of double maximum electron density peaks at the F_2 layer is known as Diurnal Double Maxima (DDM). Shun-Rung et al. (2000) attributed this structure in the middle latitudes to poleward neutral winds and thermal expansions reducing the electron concentration around noon and hence causing a midday bite-out. Detailed information about these phenomena can be found in Katamzi (2011) and Shun-Rung et al. (2000).

Figures 5.7, 5.8 and 5.9 show 2D images of the electron density on 12 July, 12 October and 21 December respectively. A diurnal trend is observed where the electron density peaks around local noon (10h00 UT). This is due to the fact that the ionosphere increases with increasing solar illumination from sunrise to roughly midday and gradually decreases afterwards. In addition it is observed that the electron density is lower in July than all the other days. This decrease in the electron concentration in the F2 layer during the winter season has also been reported by Katamzi (2008) and Habarulema et al. (2007).

Figure 5.10 demonstrates 2D electron density images at 12h00 UT for eight days for which 2 days were taken from each season of the year. The noon 2D images of the electron concentration show a seasonal variation of the peak electron concentration in which the maximum and minimum peaks are seen to be on 12 November (spring)

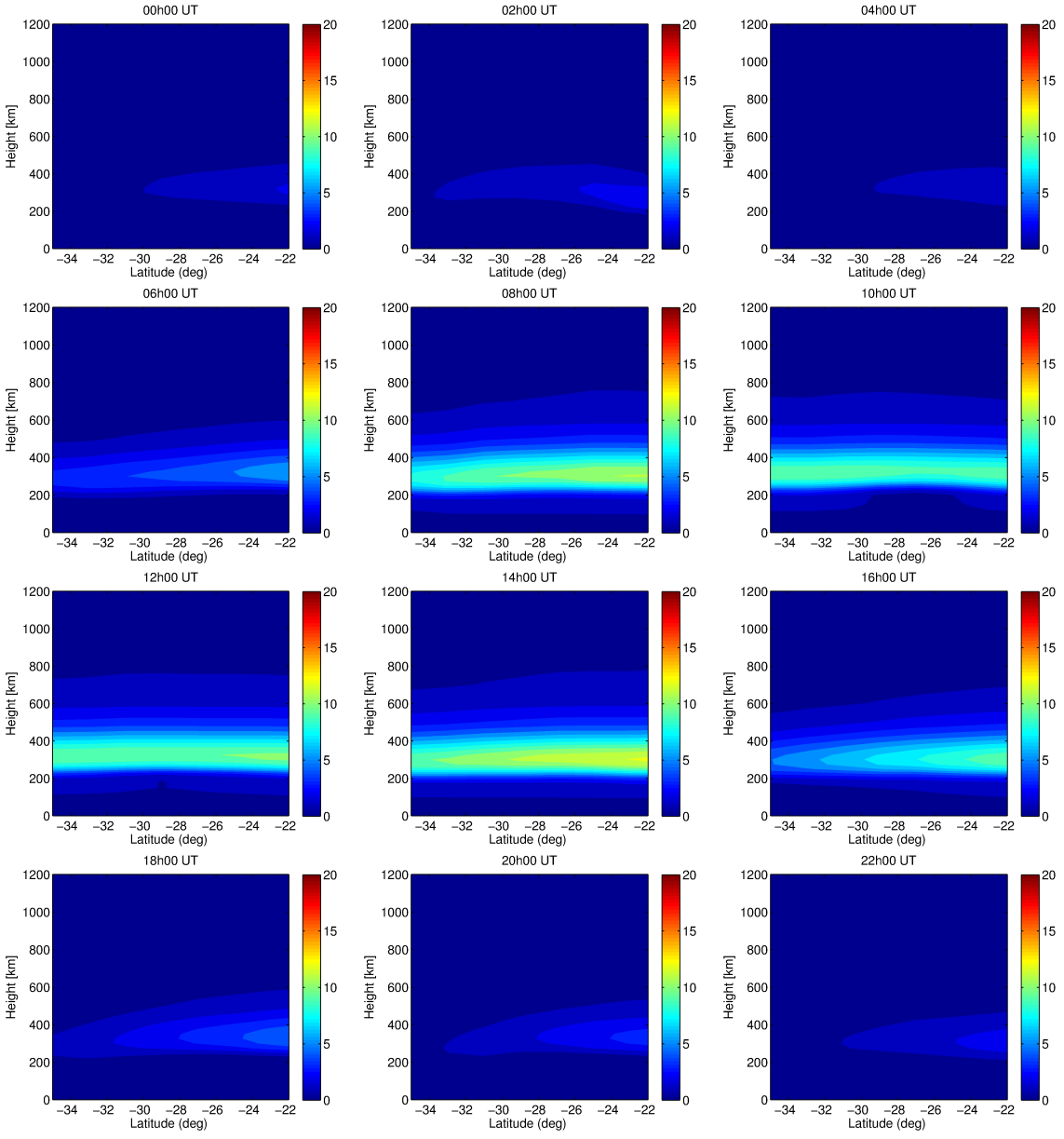

Figure 5.6: 2D electron density images from 12 April 2012

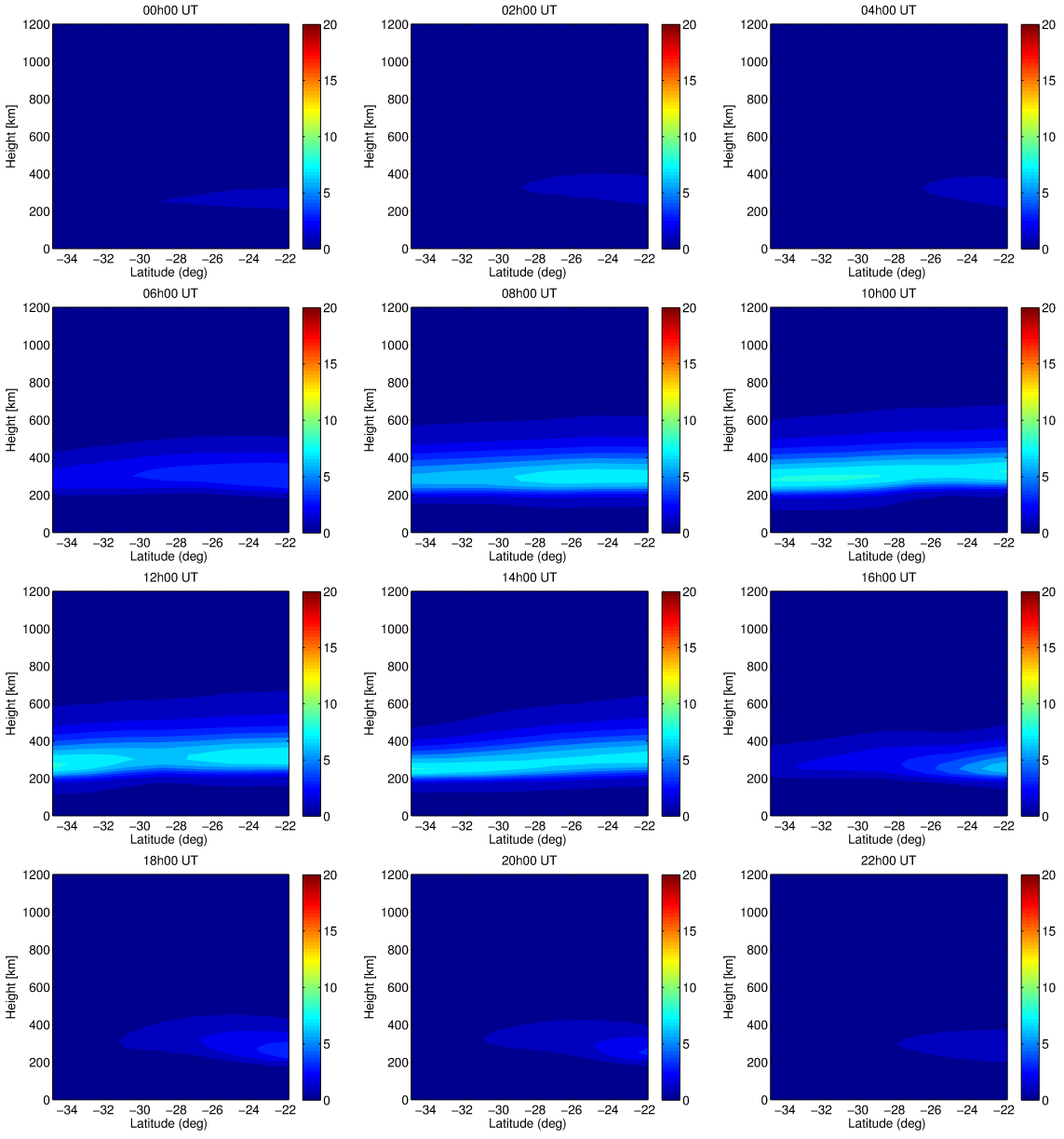

Figure 5.7: 2D electron density images from 12 July 2012

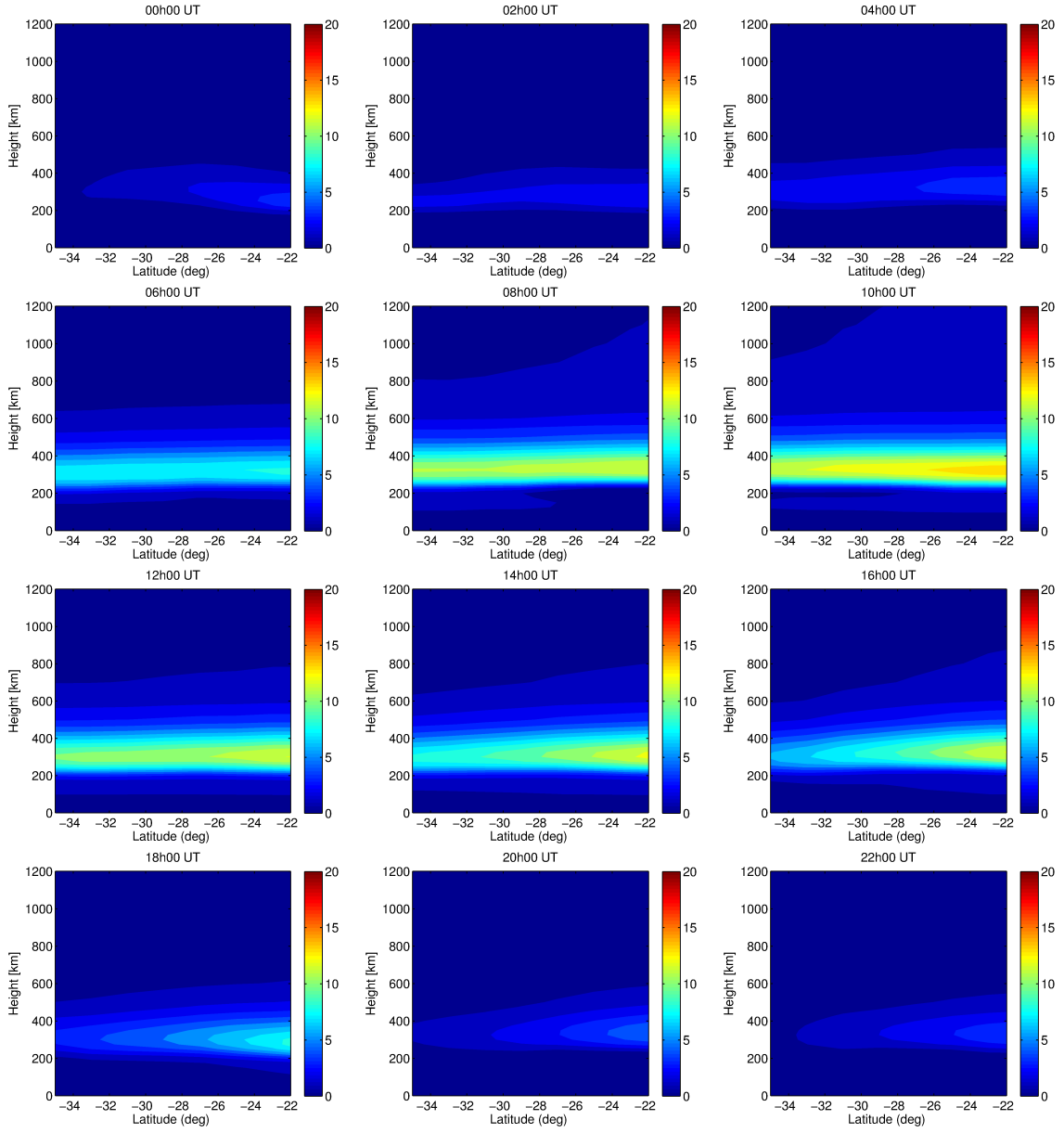

Figure 5.8: 2D electron density images from 12 October 2012

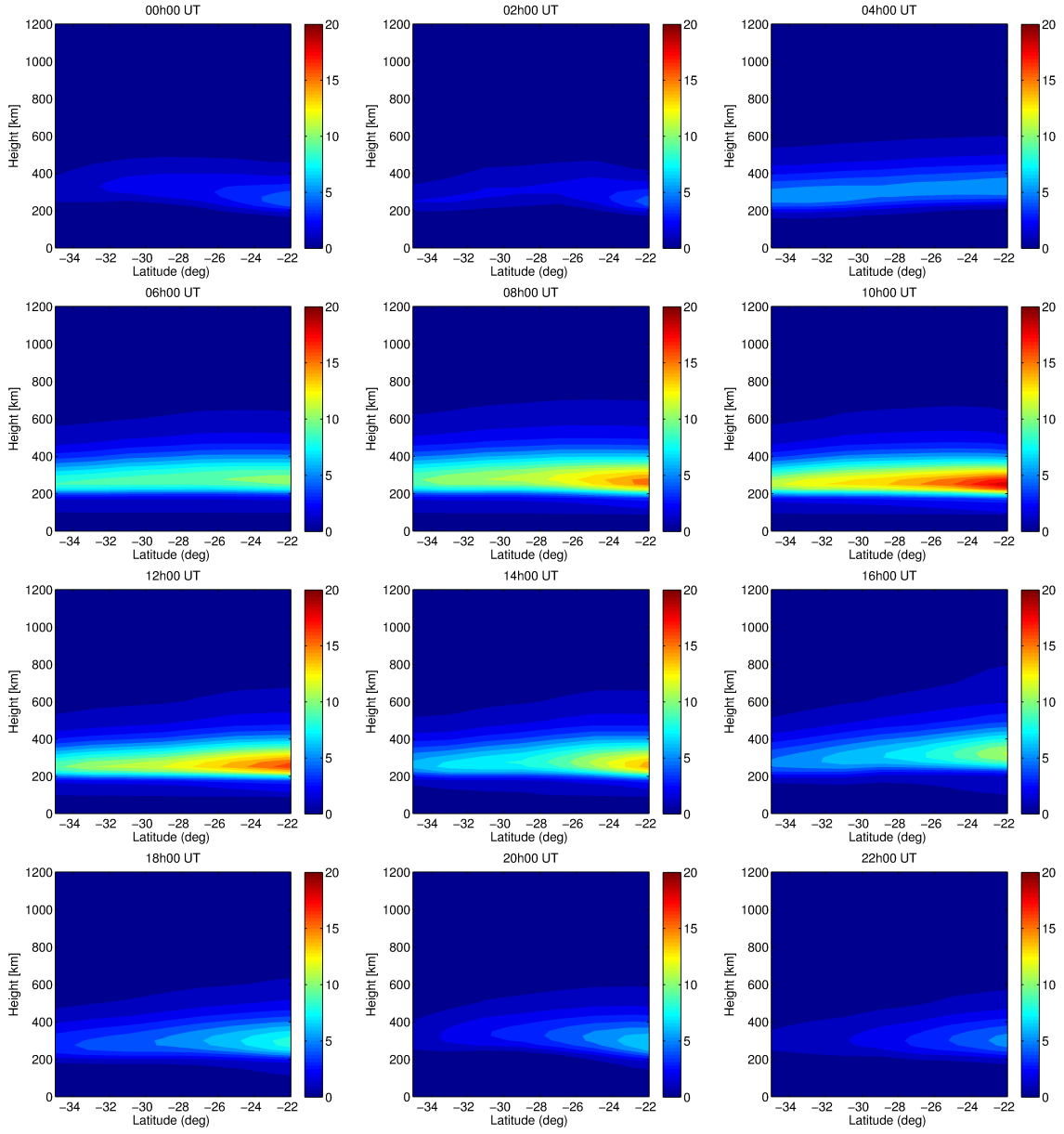

Figure 5.9: 2D electron density images from 21 December 2012

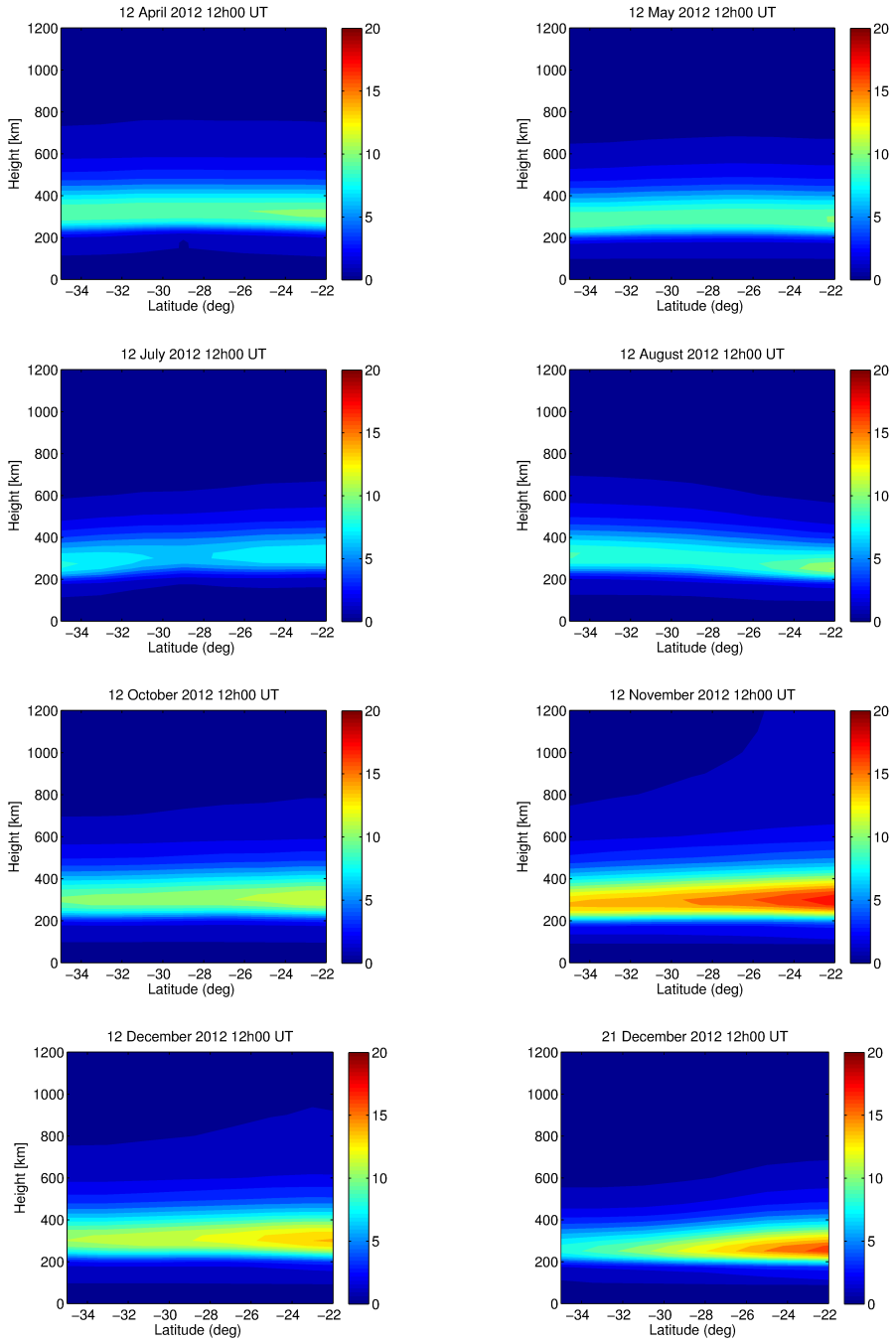

Figure 5.10: This figure shows the noon (12h00 UT) 2D electron density images of the ionosphere reconstructed by MIDAS for each validation day illustrating the seasonal variation of NmF2 for 2012.

and 12 July (winter) respectively. However studies by Habarulema (2010), Katamzi (2008) and Opperman (2008) showed that, over the South African region, the peak electron density is maximum at the equinoxes and minimum at the solstices. Since the randomly selected validation days used in this study do not fall on the solstice or equinox days, it is difficult to conclude on this difference in observations.

The noon 2D images of the peak electron densities presented in Figure 5.10 showed a seasonal variation in such a way that larger NmF2 values were observed on 12 November (spring) and on 12 and 21 December (summer) than on 12 April and 12 May (autumn) and on 12 July and 12 August (winter). The 2D images were produced at a longitude of 25.5° E which is near to the Grahamstown ionosonde station; therefore, the seasonal variation observed over the Grahamstown ionosonde is reflected on the noon 2D images of the peak electron densities. Figures 5.11

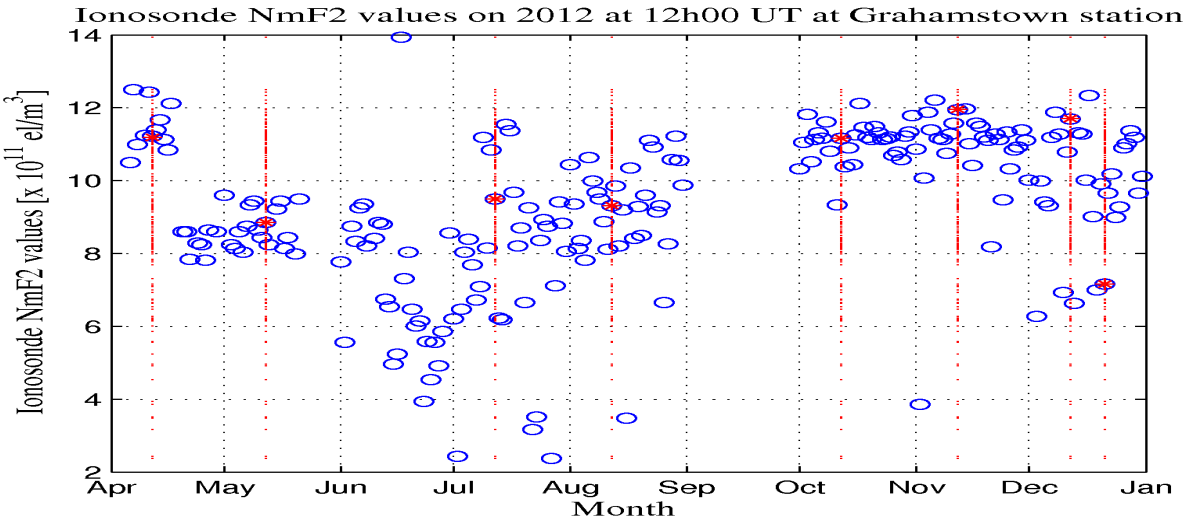

Figure 5.11: Ionosonde NmF2 values at Grahamstown ionosonde station at 12h00 UT for the year 2012. The vertical lines (red -) indicate the test dates.

and 5.12 represent the seasonal variation of ionosonde NmF2 values from the Grahamstown and Louisvale stations at 12h00 UT along with the 8 validation days indicated by red vertical lines. Even though the ionosonde NmF2 seasonal variation is more pronounced at Louisvale (refer to Figure 5.12) than at Grahamstown (refer to Figure 5.11), a seasonal trend of the NmF2 values is observed at both stations. From Figure 5.12 it is observed that maximum NmF2 values occur in April (autumn)

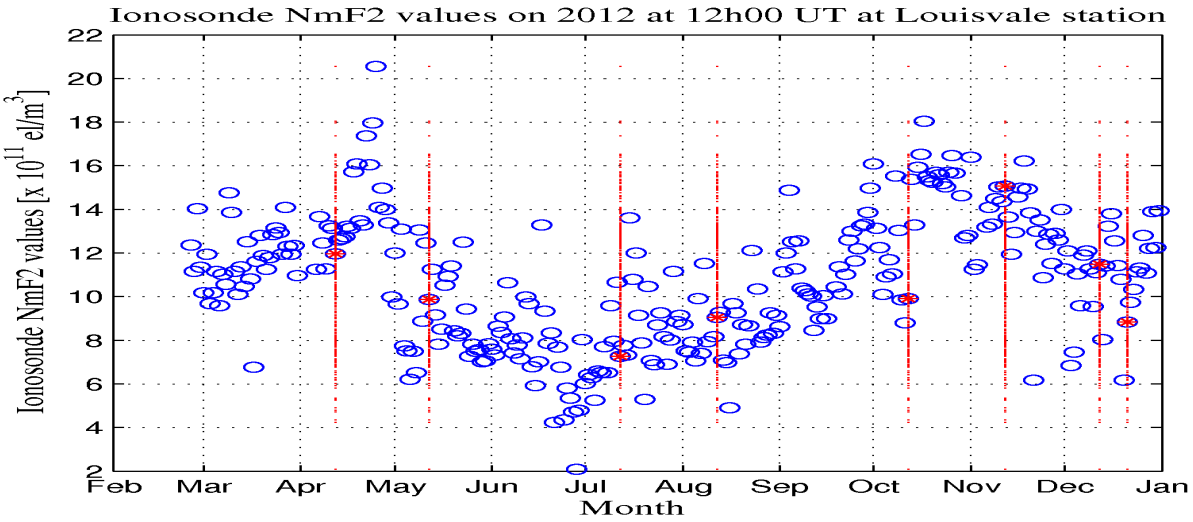

Figure 5.12: Ionosonde NmF2 values at Louisvale ionosonde station at 12h00 UT for the year 2012. The vertical lines (red -) indicate the test dates.

and October and November (spring) and minimum in July and August (winter) and December (summer) (refer to Figure 5.12); however, from the seasonal ionosonde NmF2 variation at Grahamstown station (Figure 5.11) large NmF2 values were observed in December (summer) in addition to the peak NmF2 values seen during the autumn and spring seasons.

It is observed from Figures 5.11 and 5.12 that the seasonal variation of the NmF2 values vary from station to station. Thus while the seasonal trend observations of

Habarulema (2010), Katamzi (2008) and Opperman (2008) differs from the seasonal trends in this study, it is worth noting that the seasonal trends of MIDAS are in agreement with the trend in Louisvale ionosonde measurements.

5.3 Validation

The peak electron density (NmF2) of the MIDAS reconstructed 1D electron density profile at 15 minute interval was extracted to make the comparison with a similar dimension electron density profile derived from ionosonde measurements. This was repeated for all the validation days and for all ionosonde stations whose data were available.

5.3.1 Peak Electron Density at F_2-Layer (NmF2)

12 April 2012

From Figure 5.13 (a), it is observed that the peak of the MIDAS derived NmF2 values near Grahamstown occurred before the peak of the respective ionosonde NmF2 values (i.e at about 08h00 UT). Moreover, MIDAS NmF2 values overestimated the ionosonde NmF2 values starting from 10h00 UT. However there is a good match of the MIDAS derived NmF2 values with those of the ionosonde before around 07h00 UT. The NmF2 values from MIDAS reconstructed electron densities near Hermanus underestimated the respective ionosonde NmF2 values throughout the day, refer to Figure 5.13(b). At Madimbo station the MIDAS NmF2 diurnal trend matched that of the ionosonde both showing a double NmF2 peak structure, refer to Figure 5.13 (c). The first peak occurred at around 08h00 UT for both MIDAS and ionosonde data and the second peak at around 15h00 UT for MIDAS and around 16h00 UT for the ionosonde. More information on the double peak structure, sometimes referred to diurnal double maxima (DDM) or fore-noon bite-out, can be found in the work

by Katamzi (2011) and the references therein.

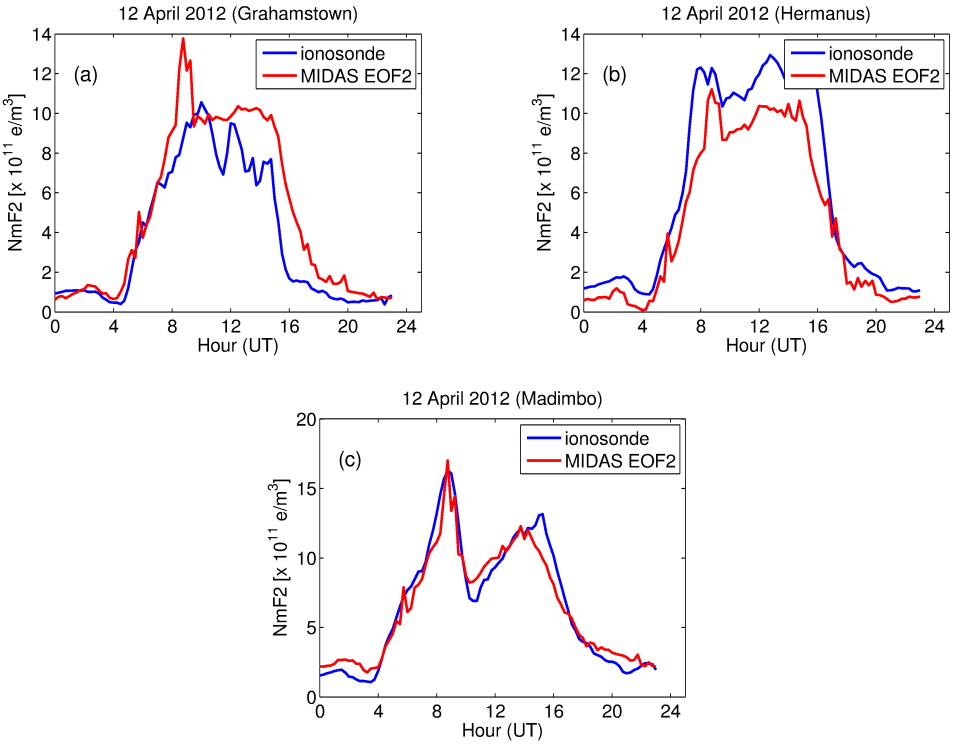

Figure 5.13: NmF2 values derived from MIDAS reconstructed electron density profiles (red -) and from ionosonde electron density profiles (blue -) at (a) Grahamstown, (b) Hermanus and (c) Madimbo on 12 April 2012.

12 May 2012

Generally MIDAS NmF2 values near Hermanus underestimated the ionosonde NmF2 values for most of the day, refer to Figure 5.14 (a). Also, it is observed that the peak of the MIDAS NmF2 values occurred a bit earlier than the ionosonde NmF2 peak. From Figure 5.14 (b) a good match of the MIDAS NmF2 and Louisvale ionosonde NmF2 trend is observed, with large deviations from the trends occurring between approximately 08h00 UT and 16h00 UT. MIDAS overestimated the

Madimbo ionosonde NmF2 values during the early and late hours of the day, refer to Figure 5.14 (c).

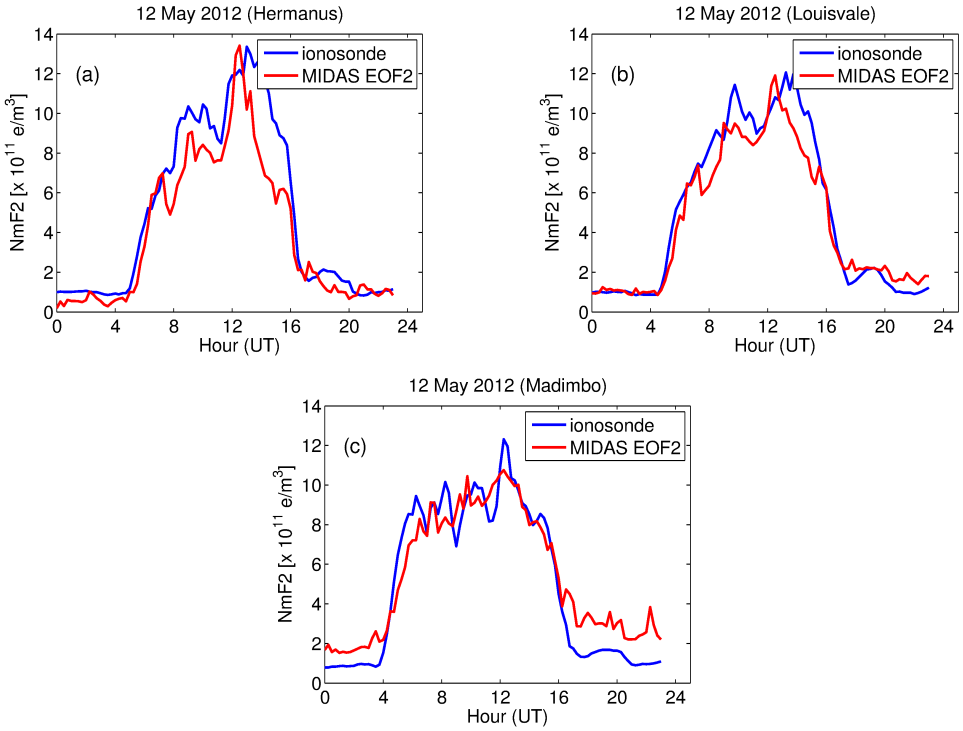

Figure 5.14: NmF2 values derived from MIDAS reconstructed electron density profiles (red -) and from ionosonde electron density profiles (blue -) at (a) Hermanus, (b) Louisvale and (c) Madimbo on 12 May 2012.

12 July 2012

Overall a good agreement is observed for the diurnal trend of the MIDAS derived NmF2 and the ionosonde NmF2 values at all four stations, refer to Figure 5.15. In particular, the trend of the reconstructed NmF2 values was better during early and late hours of the day for all the stations except at Madimbo, where MIDAS overestimated the actual NmF2 values. In addition, the reconstructed NmF2 values

showed better agreement during noon at Madimbo station than at Grahamstown, Hermanus, and Louisvale stations.

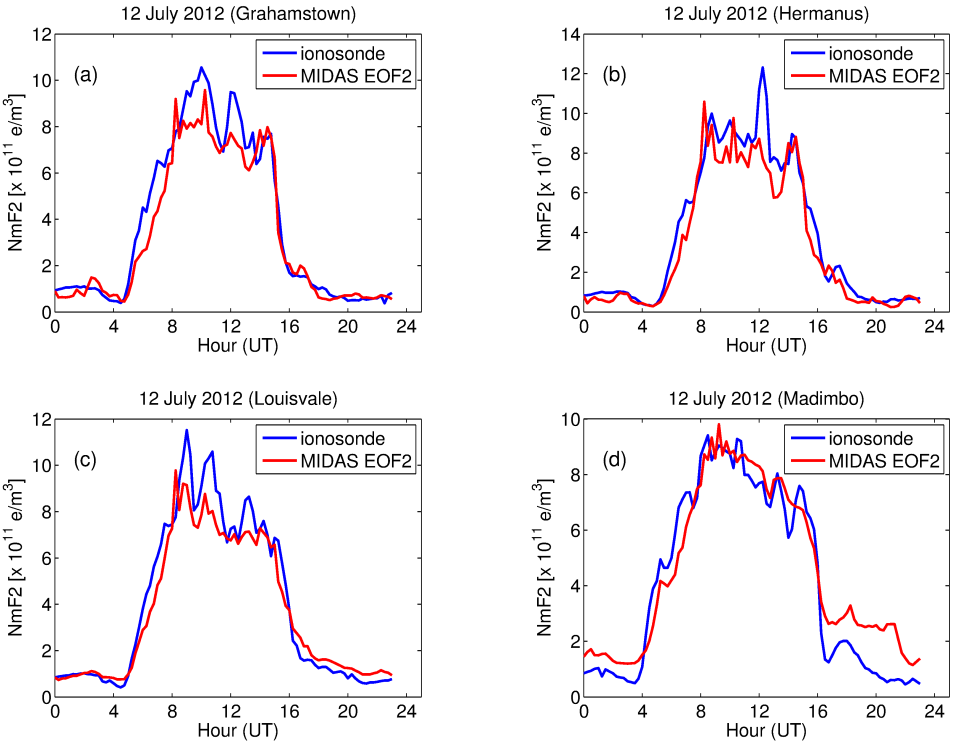

Figure 5.15: NmF2 values derived from MIDAS reconstructed electron density profiles (red -) and from ionosonde electron density profiles (blue -) at (a) Grahamstown, (b) Hermanus, (c) Louisvale and (d) Madimbo on 12 July 2012.

12 August 2012

Ionosonde data were available for only Grahamstown and Louisvale stations for this day. Figure 5.16 shows that the diurnal trend of the MIDAS derived NmF2 values fit the ionosonde NmF2 trend better over Grahamstown than over Louisvale. However, MIDAS overestimated the NmF2 values at both Grahamstown and Louisvale ionosonde stations at around 12h00 UT.

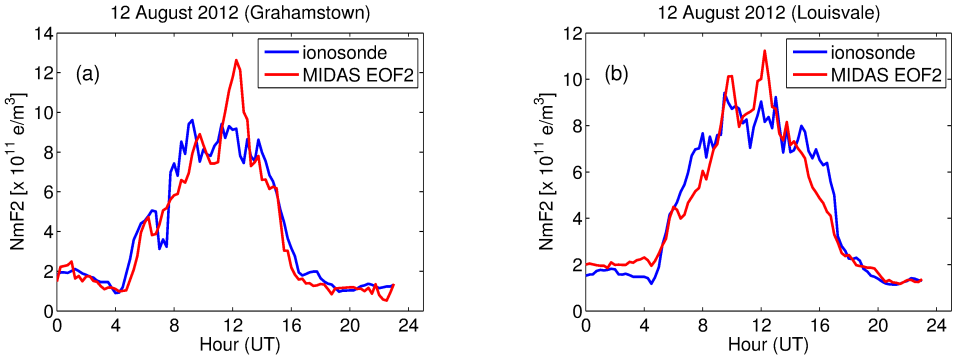

Figure 5.16: NmF2 values derived from MIDAS reconstructed electron density profiles (red -) and from ionosonde electron density profiles (blue -) at (a) Grahamstown and (b) Louisvale on 12 August 2012.

12 October 2012

NmF2 measurements were available for most of the South African ionosonde stations except for Madimbo. From Figure 5.17, it is observed that for all stations the trend of the derived NmF2 values generally matches that of the ionosonde before 08h00 UT and underestimates it for the rest of the day. In addition, MIDAS underestimated the midday peak NmF2 values for Hermanus and Louisvale but overestimated the values for Grahamstown.

12 November 2012

Ionosonde data were available for only Hermanus and Louisvale stations. There was a good match between MIDAS and ionosonde NmF2 values before 05h00 UT for both stations, as seen in Figure 5.18. However, an abrupt and short duration spike is seen in MIDAS NmF2 values around 06h00 UT for both stations (although smaller at Louisvale) and this is not visible in the ionosonde measurements. In addition, at the Hermanus station, Figure 5.18 (a), MIDAS overestimates NmF2 values between 07h00 UT and 12h30 UT and underestimates them for the rest of

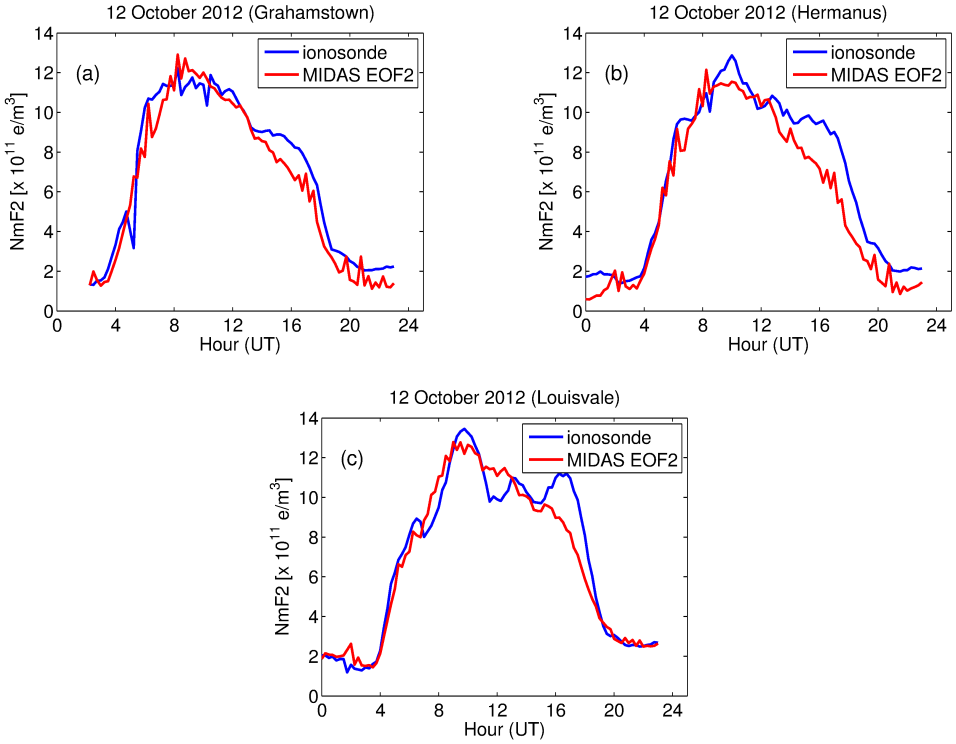

Figure 5.17: NmF2 values derived from MIDAS reconstructed electron density profiles (red -) and from ionosonde electron density profiles (blue -) at (a) Grahamstown, (b) Hermanus and (c) Louisvale on 12 October 2012.

the day. Similarly at the Louisvale station, Figure 5.18 (b), MIDAS overestimates NmF2 values between 07h00 UT and 15h30 UT and underestimates them for the rest of the day. It is also observed that NmF2 values on this day are much higher than the other validation days, except for 12 April 2012 at Madimbo.

12 December 2012

Mostly a good match of diurnal trends is seen at Grahamstown and Hermanus ionosonde stations (refer to Figures 5.19 (a) and (b) respectively), except for intermittent spikes, until around 16h00 UT. At Louisvale ionosonde station, the trends

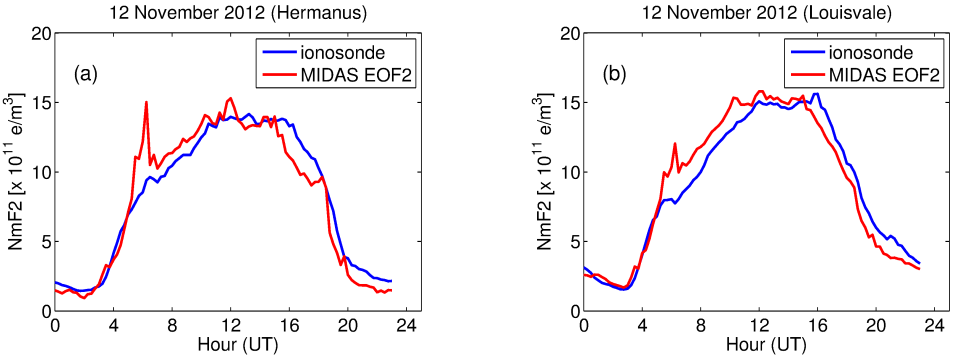

Figure 5.18: NmF2 values derived from MIDAS reconstructed electron density profiles (red -) and from ionosonde electron density profiles (blue -) at (a) Hermanus and (b) Louisvale on 12 November 2012.

matched before around 05h00 UT, and then MIDAS underestimated between 05h00 UT and 16h00 UT. Nevertheless, the reconstructed NmF2 values tend to pick up the trend exhibited by the ionosonde NmF2 values in such a way that even the slight depletion in the NmF2 at around 15h00 UT was also observed in the MIDAS reconstructed NmF2 values.

21 December 2012

Ionosonde data were available only for Grahamstown, Hermanus and Louisvale stations for this day, refer to Figure 5.20. The worst match between MIDAS and ionosonde NmF2 values is seen on this day at all three stations. The MIDAS NmF2 values showed many short term perturbations which were not exhibited by ionosonde NmF2 values. In addition, MIDAS generally overestimated ionosonde NmF2 values during the day and underestimated them in the early morning and at night. At Louisvale ionosonde station MIDAS NmF2 values were relatively smooth until around 10h00 UT. The study by Katamzi (2008) also reported that 1D MIDAS derived electron density profiles did not fit ionosonde derived ones during the summer

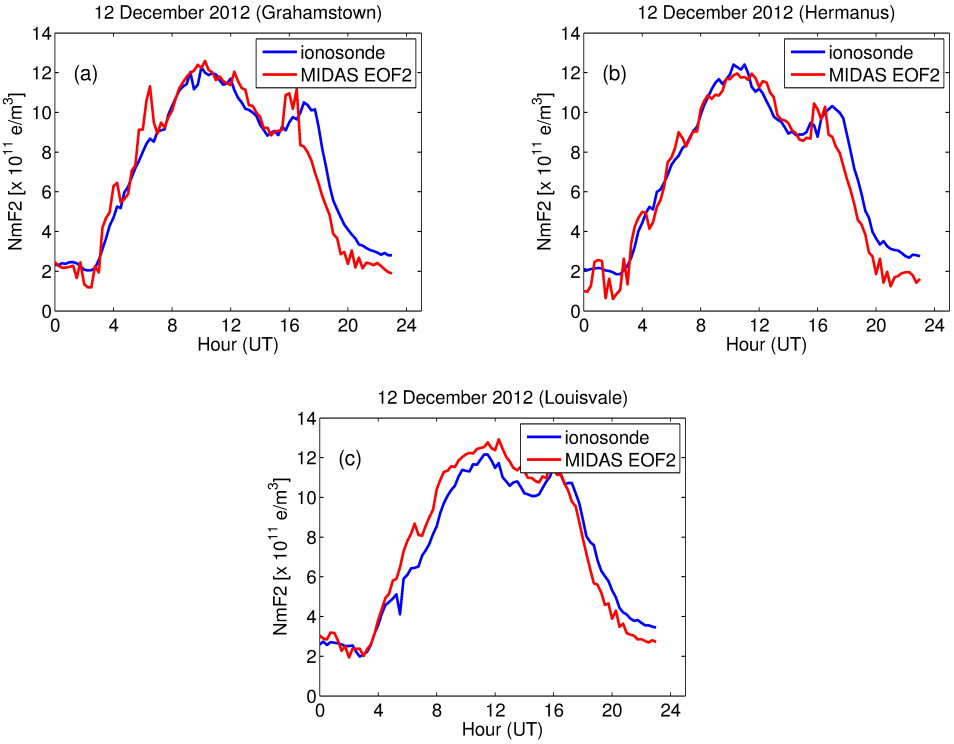

Figure 5.19: NmF2 values derived from MIDAS reconstructed electron density profiles (red -) and from ionosonde electron density profiles (blue -) at (a) Grahamstown, (b) Hermanus and (c) Louisvale on 12 December 2012.

season. Therefore both studies indicate that MIDAS underperformed in summer compared to the other seasons.

5.4 Statistical Analysis

In this section a statistical analysis, in terms of the coefficient of determination (r^2) and the root mean square error (rmse), was done in order to validate the MIDAS NmF2 results using ionosonde NmF2 measurements. The coefficient of determination (r^2) is a measure of how closely the predicted data (in this case

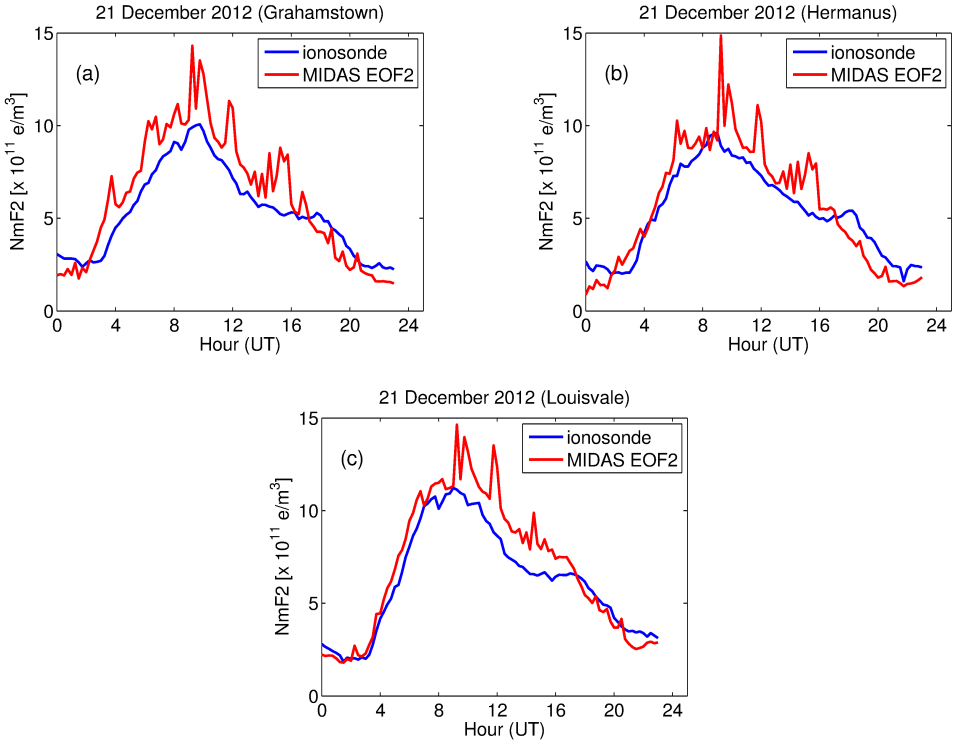

Figure 5.20: NmF2 values derived from MIDAS reconstructed electron density profiles (red -) and from ionosonde electron density profiles (blue -) at (a) Grahamstown, (b) Hermanus and (c) Louisvale on 21 December 2012.

MIDAS NmF2 values) fit the measured data (in this case ionosonde NmF2 values). The value of r^2 ranges between 0 and 1. Therefore, an r^2 value closer to one means a greater proportion of variance of the measured ionosonde values is accounted for by the predicted/reconstructed NmF2 values. Simply put, r^2 is a measure of how much of the variation in the ionosonde NmF2 value is seen by MIDAS NmF2. The r^2 showed in Figures 5.21 - 5.28 were generally greater than 0.90 but less than 1 for all stations for all validation days.

However, r^2 values alone do not provide much information about how accurate the predicted/reconstructed outputs (i.e MIDAS NmF2 values) are compared with

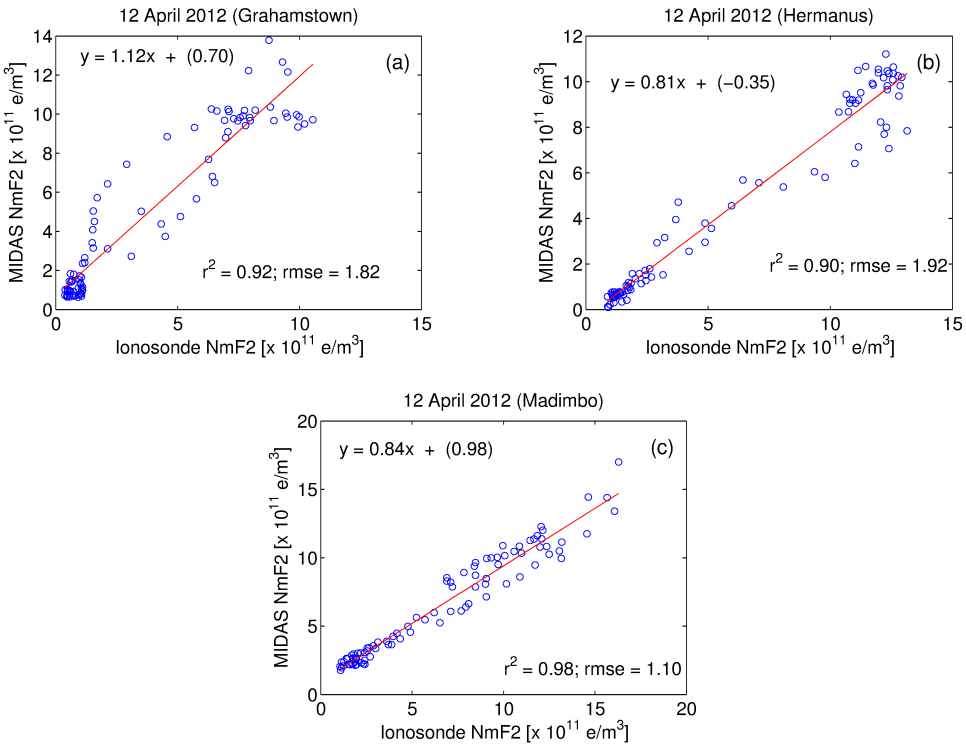

Figure 5.21: Scatter plots used to compare MIDAS NmF2 values with ionosonde NmF2 values from (a) Grahamstown, (b) Hermanus and (c) Louisvale on 12 April 2012.

the measured NmF2 values (i.e ionosonde NmF2 values) hence a root mean square error (rmse) between ionosonde NmF2 values and MIDAS NmF2 values was also calculated. The rmse is a measure of the spread of the measured values (in this case ionosonde NmF2 values) about the predicted/reconstructed values (in this case MIDAS NmF2 values). In this study, the rmse value is defined as:

$$\text{rmse} = \sqrt{\frac{\sum_{i=1}^{N}(\text{MIDAS}_{\text{NmF}_2} - \text{Ionosonde}_{\text{NmF}_2})^2}{N}} \tag{5.1}$$

Overall MIDAS underestimated the NmF2 values from the ionosondes on 12 July (at all stations), on 12 May (at Hermanus, Louisvale and Madimbo), on 12 April (at

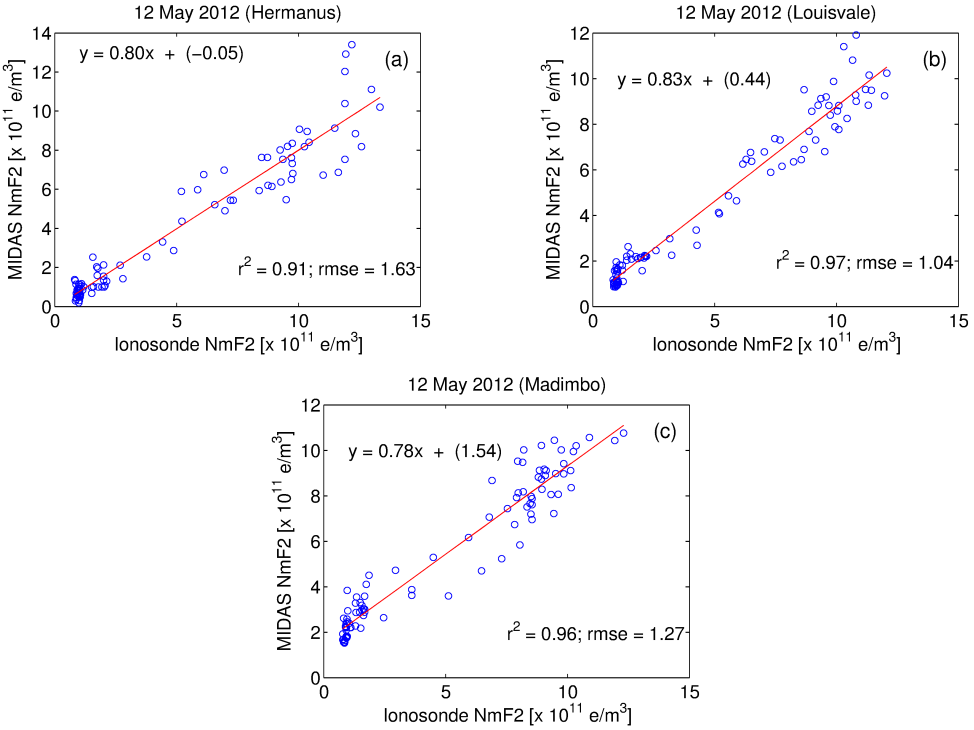

Figure 5.22: Scatter plots used to compare MIDAS NmF2 values with ionosonde NmF2 values from (a) Hermanus, (b) Louisvale and (c) Madimbo on 12 May 2012.

Hermanus and Madimbo), 12 August (at Louisvale) and on 12 October (at Hermanus and Louisvale stations). This observation is supported by the fact that the slopes are greater than 1 on those days and at those stations (refer to Figures 5.23, 5.24, 5.25, 5.21 and 5.22 respectively). Meanwhile, MIDAS generally overestimated NmF2 values in November (at Hermanus and Louisvale), in December (at Grahamstown, Hermanus and Louisvale) and in April, August and October (at Grahamstown station); refer to Figures 5.26, 5.27, 5.28, 5.21, 5.24 and 5.25 respectively. In general, MIDAS overestimated the NmF2 values at Grahamstown station on all validation days except 12 July.

The rmse values between the MIDAS NmF2 and ionosonde derived NmF2 values

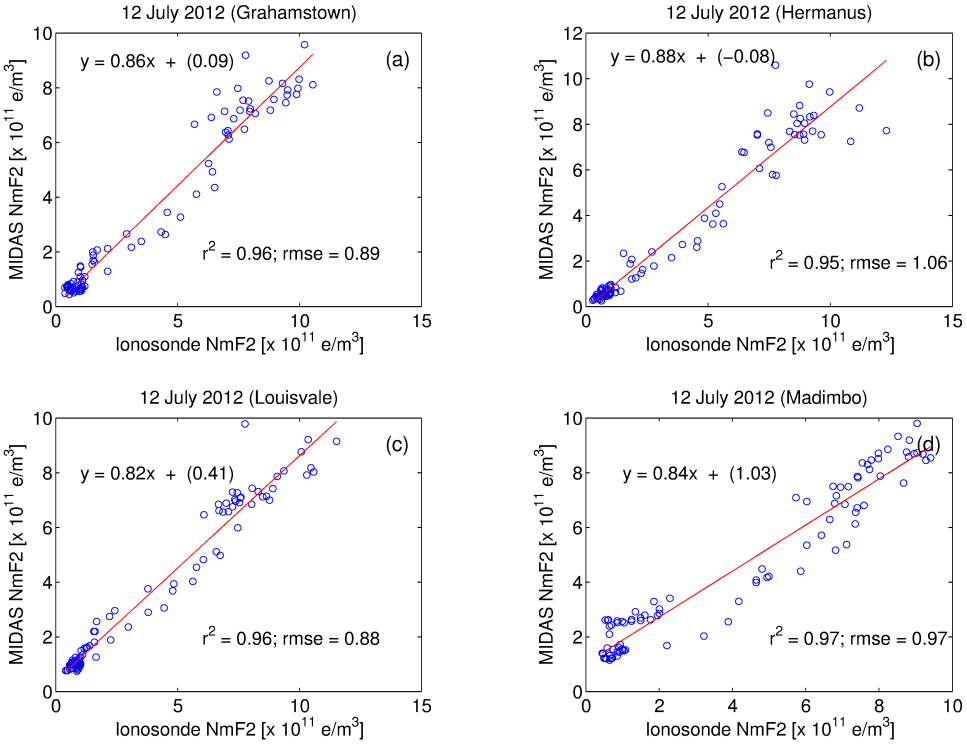

Figure 5.23: Scatter plots used to compare MIDAS NmF2 values with ionosonde NmF2 values from (a) Grahamstown, (b) Hermanus, (c) Louisvale and (d) Madimbo on 12 July 2012.

were the least at three of the South African ionosonde stations on 12 July 2012, except for Hermanus station for which the rmse value was least on 12 December. In fact the least rmse value (i.e 0.88x $10^{11}[el/m^3]$) from all ionosonde stations and all validation days was obtained on 12 July 2012 at Louisvale ionosonde station. This indicates that the MIDAS reconstructed NmF2 values during the winter season have minimum rmse values compared to the other seasons of the year. A maximum rmse values of 1.92x $10^{11}[el/m^3]$ and 1.82x $10^{11}[el/m^3]$ (refer to Figures 5.21 (a) and (b)) were obtained at Hermanus and Grahamstown stations respectively on 12 April. While the maximum rmse value over Louisvale was 1.32x $10^{11}[el/m^3]$ on 21

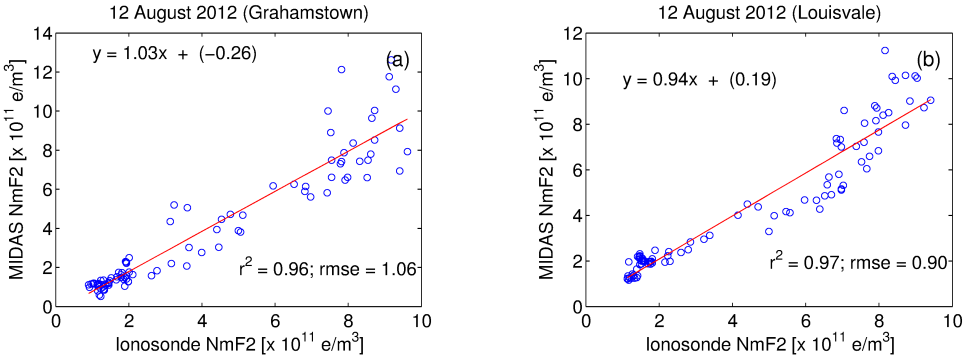

Figure 5.24: Scatter plots used to compare MIDAS NmF2 values with ionosonde NmF2 values from (a) Grahamstown and (b) Louisvale on 12 August 2012.

December 2012 and over Madimbo this value was 1.27x $10^{11}[el/m^3]$ on 12 May 2012. This indicates that in general the MIDAS NmF2 values have the highest rmse values in autumn than the other seasons at most stations.

Even though the rmse values on 21 December were large compared to the rmse values on 12 December (both summer days), the average rmse value in summer was still less than the average rmse value in autumn (12 April and 12 May) and spring (12 October and 12 November). In fact the averaged rmse values at all stations in winter (12 July and 12 August), summer (12 December and 21 December), autumn (12 April and 12 May) and spring (12 October and 12 November) were 0.96x, 1.24x, 1.34x, and 1.46x $10^{11}[el/m^3]$ respectively. This demonstrates that on a seasonal basis, on average MIDAS reconstructed NmF2 values fit relatively better in winter and summer than in spring and autumn seasons of the year. The MIDAS produced NmF2 values for the 21 December were highly overestimated and showed intermittent spikes that were not observed in the ionosonde NmF2 values (refer to Figure 5.28 (a) to (c)); and therefore, can not be trusted to draw conclusions.

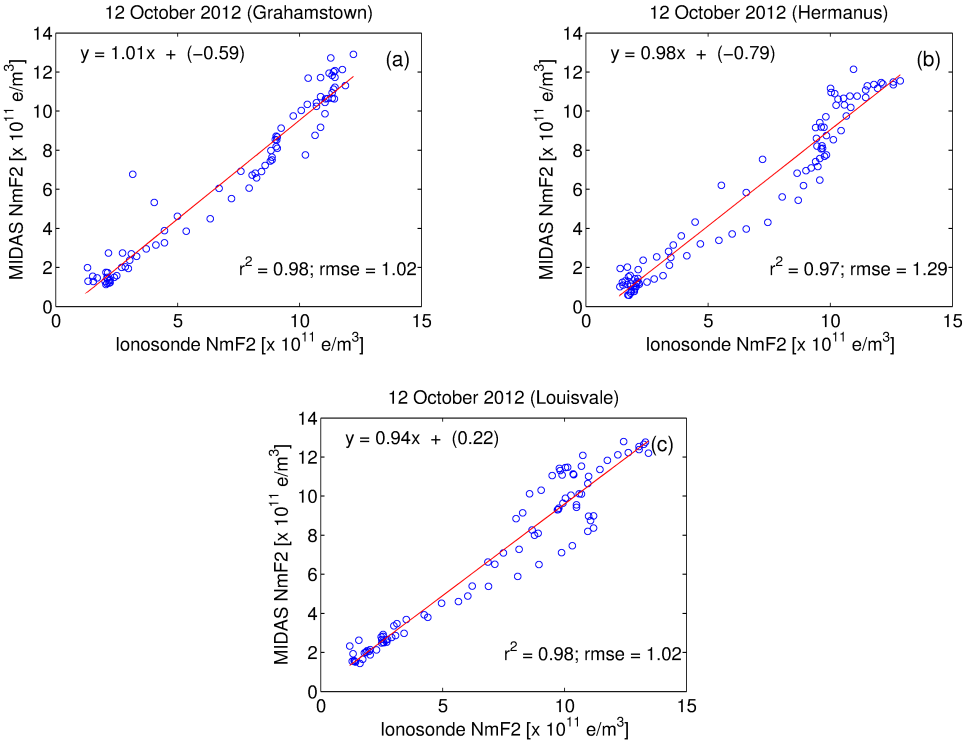

Figure 5.25: Scatter plots used to compare MIDAS NmF2 values with ionosonde NmF2 values from (a) Grahamstown, (b) Hermanus and (c) Louisvale on 12 October 2012.

The rmse values were also varying from station to station within the days of inversion. Table 5.3 summarizes the maximum and minimum rmse values calculated between the reconstructed and measured values for all stations. This indicates that the rmse values between the MIDAS reconstructed and ionosonde derived NmF2 values vary significantly, the maximum being during autumn at Grahamstown, Hermanus and Madimbo stations and the minimum during the winter season for all but Hermanus stations (refer to Table 5.3). In fact the rmse values are larger at Grahamstown and Hermanus stations than at Louisvale and Madimbo (refer to Figures 5.21, 5.22 and 5.28 and Table 5.3).

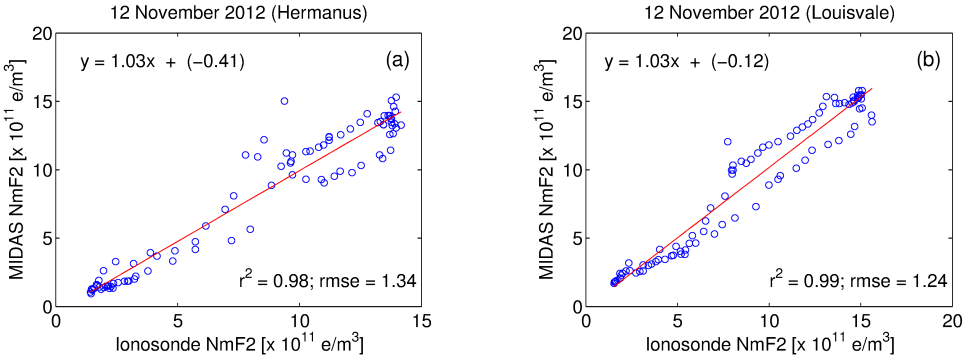

Figure 5.26: Scatter plots used to compare MIDAS NmF2 values with ionosonde NmF2 values from (a) Hermanus and (b) Louisvale on 12 November 2012.

Table 5.3: A summary of the maximum and minimum rmse values calculated at all four stations and number of days for which ionosonde data was available (out of the 8 Validation days) along with the days in which the values were obtained.

		RMSE [x $10^{11} el/m^3$]			
Station Name	No.of Validation days	Max	Val. day	Min	Val. day
Grahamstown	6	1.82	12 April	0.89	12 July
Hermanus	7	1.92	12 April	1.01	12 December
Louisvale	7	1.32	21 December	0.88	12 July
Madimbo	3	1.27	12 May	0.97	12 July

Apart from the rmse values, the r^2 values also show variation from station to station and season to season. For instance, the lowest r^2 value for Hermanus was 0.90 on 12 April and 12 May and the highest 0.98 on 12 November and 12 December, also the highest r^2 values for Grahamstown was 0.98 on 12 October and 12 December and the lowest was 0.92 on 12 April. Meanwhile, in Louisvale the lowest value was 0.96 in July and August and highest was 0.99 in November and December. Lastly, for Madimbo station the lowest r^2 value was 0.96 in May and highest 0.99 in April. This indicates that at most stations generally the lowest r^2 value occurred during autumn and the highest during summer. Also that MIDAS performs better, in terms of r^2, at Louisvale and Madimbo than at Hermanus and Grahamstown.

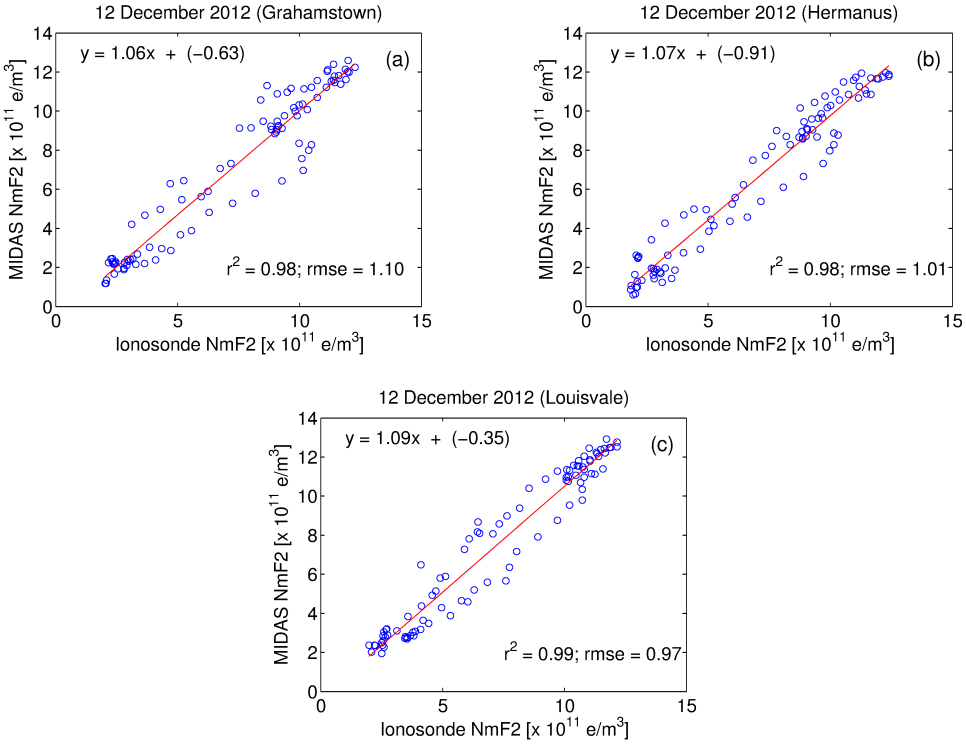

Figure 5.27: Scatter plots used to compare MIDAS NmF2 values with ionosonde NmF2 values from (a) Grahamstown, (b) Hermanus and (c) Louisvale on 12 December 2012.

In summary, this analysis confirmed that the highest rmse values were calculated during autumn while the r^2 were lowest in autumn and highest in summer. In addition, NmF2 values over Hermanus had the worst minimum and maximum rmse values and the lowest r^2 value than the other ionosonde stations while Madimbo had the lowest rmse value and one of the highest r^2 values (refer to Figures 5.23 - 5.22 and Table 5.3).

In conclusion, MIDAS is observed to perform better during winter and summer than spring and autumn. A study by Katamzi (2008) also indicated that MIDAS1 (the first version) produced noon NmF2 values that reproduced measured NmF2 values

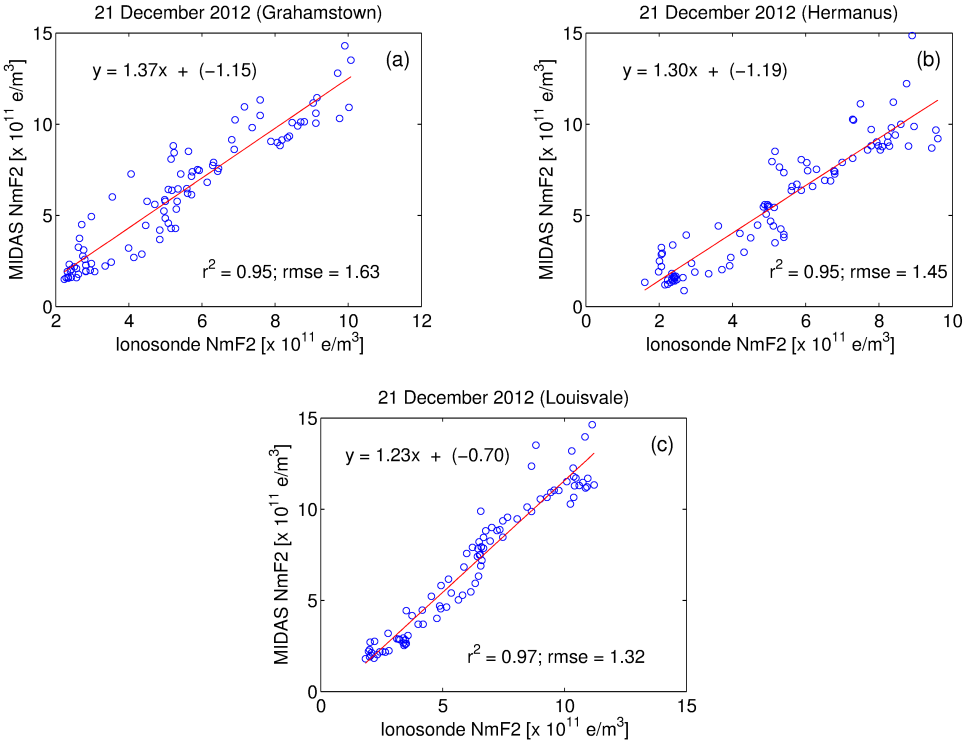

Figure 5.28: Scatter plots used to compare MIDAS NmF2 values with ionosonde NmF2 values from (a) Grahamstown, (b) Hermanus and (c) Louisvale on 21 December 2012.

with greater accuracy during the winter season than during other seasons.

Chapter 6

Conclusion and Future Work

6.1 Discussion and Conclusion

The main aim of the research presented in this thesis was to optimize the MIDAS algorithm in order to make available MIDAS reconstructed electron density profiles over the South African region for the study of different ionospheric phenomena. The reliability of MIDAS reconstructions depends mostly on the choice of the method used to construct EOFs (choice of Chapman or IRI95 model), the number of EOFs used in the inversion process (between 1 and 5) and the inversion method (choice of *minres* or *quadprog*). Through a process of elimination (Section 5.1) 2 EOFs constructed from the IRI95 model and the *quadprog* method were found to be the optimum parameters, used to produce the results presented.

2D electron density maps from the MIDAS inversion procedure were presented for different days in the year 2012 to demonstrate seasonal and geographic variation of the ionosphere over the South African region. In order to validate the MIDAS reconstructions, the MIDAS and ionosonde NmF2 values extracted from their respective 1D electron density profiles at 15 minute intervals were compared and the rmse and r^2 values computed in order to evaluate MIDAS's performance.

The comparison of the MIDAS reconstructed NmF2 and ionosonde NmF2 values showed that the least rmse values (i.e 0.88x, 0,89x, and 0.97x $10^{11}[el/m^3]$) were found to occur on 12 July at Louisvale, Grahamstown and Madimbo ionosonde stations respectively (see Table 5.3). This indicated that MIDAS performs more accurately during the winter than during the other seasons. In addition, maximum rmse values of 1.92x $10^{11}[el/m^3]$ and 1.82 x $10^{11}[el/m^3]$ were computed between the MIDAS NmF2 and ionosonde NmF2 values on 12 April at Hermanus and Grahamstown ionosonde stations respectively; therefore, this indicated that MIDAS reconstruction was associated with large errors during the autumn season. A rmse value of 1.63x $10^{11}[el/m^3]$ over the Hermanus ionosonde station on 12 May was another indication of large errors found during the autumn season. A minimum r^2 value of 0.90 on 12 April and 12 May (both at Hermanus) and a r^2 value of 0.92 observed over the Grahamstown station on 12 April also strengthens the analysis that MIDAS is associated with large errors during the autumn. Meanwhile a maximum r^2 value of 0.99 was obtained in November and December for the Grahamstown station and the r^2 values were between 0.95 and 0.97 for all the stations in July indicating that, in general, MIDAS reproduces the measured data more accurately, in terms of r^2 values, during the summer and winter.

Overall MIDAS reconstructed NmF2 values reproduced their respective ionosonde NmF2 values more accurately during the winter and summer than during the autumn and spring seasons (refer to Figures 5.23 - 5.22) except on 21 December where the MIDAS reconstructed NmF2 values exhibited many perturbations which were not observed in the ionosonde NmF2 values. Also MIDAS performs better, in terms of r^2, at Louisvale and Madimbo stations than Hermanus and Grahamstown stations.

In conclusion, the daily basis comparison between the MIDAS reconstructed NmF2

and ionosonde derived NmF2 values showed good agreement. This indicates that the reconstruction of the electron density profiles by the MIDAS algorithm using GPS data from the South African GPS receivers is useful for 3D mapping of the ionosphere over this region as well as the study of ionospheric phenomena.

6.2 Future Work

The MIDAS inversion used data from approximately 53 GPS receivers and their distribution was nonuniform which impacts the reconstructed profiles, as discussed by Zapfe et al. (2006). Therefore, the MIDAS reconstructed electron density profiles would provide a more accurate representation of the measured data if GPS data from a denser and more evenly distributed network was available. It is also important to assess MIDAS under all geomagnetic conditions and solar activities, so that a near real-time imaging of the ionosphere can be made possible for the South African region.

Since the time dependent 3D inversion using MIDAS have picked the physical ionospheric phenomena, such as the midday bite-outs at the F_2 layer of the ionosphere, future work would include the use of MIDAS to study large scale regions of depleted equatorial ionospheric plasma called equatorial bubbles.

References

Allain, D. and Mitchell, C. N.: Ionospheric delay corrections for single-frequency GPS receivers over Europe using tomographic mapping, GPS Solutions, 13, 141–151, 2009.

Austen, J., Franke, S., Liu, C., and Yeh, K.: Application of Computerized tomography techniques to ionospheric research, pp. 25–35, International Beacon Satellite Syposium; Oulu, Finland, 1986.

Austen, J. R., Franke, S. J., and Liu, C. H.: Ionospheric imaging using computerized tomography, Radio Science, 23, 299–307, 1988.

Bauer, S.: On the structure of the topside ionosphere, Journal of the atmospheric sciences, 16, 276–278, 1962.

Baumjohann, W. and Treumann, R.: Basic space plasma physics, Imperial college Press, London, 1997.

Bilitza, D.: International Reference Ionosphere 2000, Radio Science, 36, 261–275, 2001.

Bracewell, R.: Strip Integration in Radio Astronomy, Australian Journal of Physics, 9, 198, 1956.

Bust, G. S. and Mitchell, C. N.: Review of the Current Status of Four-Dimensional Ionospheric Imaging, in: In Characterising the Ionosphere ., vol. 46 of *Meeting Proceedings RTO-MP-IST-056, Paper 31*, pp. 31–1–31–18, Neuilly-sur-Seine, France: RTO., 2006.

Bust, G. S. and Mitchell, C. N.: History, current state, and future directions of ionospheric imaging, Reviews of Geophysics, 46, 2008.

Chartier, A. T., Mitchell, C. N., and Jackson, D. R.: A 12 year comparison of MIDAS and IRI 2007 ionospheric Total Electron Content, Advances in Space Research, 49, 1348–1355, 2012.

Chen, F.: Introduction to Plasma Physics and Controlled Fusion, V. 1, Springer, Los Angeles, 1984.

Christie, J. R., Parkinson, B. W., and Enge, P. K.: The effects of the ionosphere and C/A frequency on GPS signal shape: considerations for GNSS-2, INST of Navigation, Alexandria, VA,(USA)., 1, 647–653, 1996.

Cilliers, P. J., Opperman, B. D., Mitchell, C. N., and Spencer, P. J.: Electron density profiles determined fromtomographic reconstruction of total electron content obtained from GPS dual frequency data: First results from the South African network of dual frequency GPS receiver stations, Advances in Space Research, 34, 2049–2055, 2004.

Davies, K.: Ionospheric radio, IEEE Electromagnetic Waves Series, Peter Peregrinus, London, 1990.

Dear, R. M. and Mitchell, C. N.: Ionospheric imaging at mid-latitudes using both GPS and ionosondes, Journal of Atmospheric and Solar-Terrestrial Physics, 69, 817–825, 2007.

Ercha, A., Zhang, D., Ridley, A. J., Xiao, Z., and Hao, Y.: A global model: Empirical orthogonal function analysis of total electron content 1999–2009 data, Journal of Geophysical Research, 117, 2012.

Farrell, J. and Barth, M.: The Global Positioning System & Inertial Navigation, McGraw-Hill Companies,Incorporated, New York, 1998.

Fehmers, G.: Tomography of the Ionosphere, Technische Universiteit Eindhoven, 1996.

Fremouw, E. J., Secan, J. A., and Howe, B. M.: Application of stochastic inverse theory to ionospheric tomography, Radio Science, 27, 721–732, 1992.

Habarulema, J.: A contribution to TEC modelling over Southern Africa using GPS data, Ph.D. thesis, Rhodes University, Grahamstown, South Africa, 2010.

Habarulema, J., Mckinnell, L.-A., and Cilliers, P.: Prediction of global positioning system total electron content using Neural Networks in South Africa, Journal of Atmospheric and Solar-Terrestrial physics, 69, 1842–1850, 2007.

Hajj, G. A., Wilson, B. D., Wang, C., Pi, X., and Rosen, I. G.: Data assimilation of ground GPS total electron content into a physics-based ionospheric model by use of the Kalman filter, Radio Science, 39, 2004.

Hofmann-Wellenhof, B., Lichtenegger, H., and Collins, J.: Global Positioning System: theory and practice, Springer-Verlag, Wien, New York, 4th edition, 1997.

Hoque, M. M. and Jakowski, N.: Ionospheric Propagation Effects on GNSS Signals and New Correction Approaches, 2012.

Howe, B. M., Runciman, K., and Secan, J. A.: Tomography of the ionosphere: Four-dimensional simulations, Radio Science, 33, 109–128, 1998.

Katamzi, Z.: Statistical Analysis of Ionospheric Total Electron Content, Ph.D. thesis, University of Bath, 2011.

Katamzi, Z. T.: Verification of Ionospheric tomography using MIDAS over Grahamstown, South Africa, Master's thesis, Rhodes University, 2008.

Kunitsyn, V. and Tereshchenko, E.: Ionospheric Tomography, Physics of Earth and Space environment, Springer-Verlag, Berlin, 2003.

Liu, J. Y., Chen, Y. I., Jhuang, H. K., and Lin, Y. H.: Ionospheric foF2 and TEC Anomalous Days Associated with M $\geq$ 5.0 Earthquakes in Taiwan during 1997-1999, TAO, 15, 371–383, 2004.

Loewe, C. A. and Prolss, G. W.: Classification and mean behaviour of magnetic storms, Journal of Geophysical Research, 102, 14 209–14 562, 1997.

Mannucci, A., Iijima, B., Lindqwister, U., Pi, X., Sparks, L., and Wilson, B.: GPS and Ionosphere : Revised edition to URSI reviews of Radio Science, 1999.

Materassi, M. and Mitchell, C. N.: Imaging of the equatorial ionosphere, Annals of Geophysics, 48, 477–482, 2005.

McNamara, L.: The ionosphere: communications, surveillance, and direction finding, Orbit, a foundation series, Krieger Pub. Co., 1991.

Meggs, R. W., Mitchell, C., Watson, R. J., and R.M., D.: Real-Time Imaging of the Ionosphere over the United Kingdom – Preliminary Results. In Characterising the Ionosphere, pp. 27–1–27–6, Meeting Proceedings RTO-MP-IST-056 , Paper 27. Neuilly-sur-Seine, France, Neuilly-sur-Seine, France, 2006.

Misra, P. and Enge, P.: Global positioning system: signals, measurements, and performance, Ganga-Jamuna Press, Massachusetts, 2006.

Mitchell, C. N. and Spencer, P. S. J.: A three-dimensional time-dependent algorithm for ionospheric imaging using GPS, Annals of Geophysics, 46, 687–696, 2003.

Muella, M. T. A. H., de Paula, E. R., Mitchell, C. N., Kintner, P. M., Paes, R. R., and Batista, I. S.: Tomographic imaging of the equatorial and low-latitude ionosphere over central-eastern Brazil, Earth Planets and Space, 63, 129–138, 2011.

NAVSTARGPS: NAVSTARGPS , Navigation Signal Timing and Ranging (NAVSTAR) GPS user eqipment introduction (Public Release Version), URL `www.navcen.uscg.gov/pubs/gps/gpsuser/gpsuser.pdf`, 1996.

Okoh, D. I.: Developing an ionospheric map for South Africa, Master's thesis, Rhodes University, 2009.

Opperman, B. D. L.: Reconstructing ionospheric TEC over South Africa using signals from a regional GPS network, Ph.D. thesis, Rhodes University, 2008.

Parkinson, B. and Spilker, J.: Global Positioning System: Theory and Applications (volume One), Progress in astronautics and aeronautics, American Institute of Aeronautics and Astronautics, Inc., The University of Michigan, 5 edn., 1996.

Reinisch, B. W., Huang, X. Q., Belehaki, A., Shi, J. K., Zhang, M. L., and Ilma, R.: Modeling the IRI topside profile using scale heights from ground-based ionosonde measurements, Advances in Space Research, 34, 2026–2031, 2004.

Rishbeth, H. and Garriott, O.: Introduction to ionospheric physics, vol. 14 of *International geophysics series*, Academic Press, New York, 1969.

Rose, J.: Ionospheric Imaging to Improve GPS Timing, Ph.D. thesis, University of Bath, 2011.

Selcher, C.: Three-dimensional, High Resolution, Computerized Ionospheric Tomographic Imaging and Computational Modeling of an Artificial Ionospheric Cavity, West Virginia University, 2007.

Shun-Rung, Z., Oliver W., L., Shoichiro, F., and Yuichi, O.: A study of the forenoon ionospheric F2 layer behavior over the middle and upper atmospheric radar, Journal of Geophysical Research, 105, 15,823–15,833, 2000.

Sibanda, P.: Challenges in Topside Ionospheric Modelling Over South Africa, Ph.D. thesis, Rhodes University, 2010.

Spencer, P. S. J.: MIDAS, version3, Tech. rep., University of Bath, Bath, 2008.

Spencer, P. S. J. and Mitchell, C. N.: Imaging of fast moving electron-density structures in the polar cap, Annals of Geophysics, 50, 427–434, 2007.

Van Allen, J.: Origins Of Magnetospheric Physics: An Expanded Edition, University of Iowa Press, Iowa City, 2004.

Yeh, K. and Raymund, T.: Limitations of ionospheric imaging by tomography, Radio Science, 26, 1361–1380, 1991.

Yizengaw, E.: Imaging the ionosphere, Ph.D. thesis, La Trobe University, 2004.

Yizengaw, E., Dyson, P. L., Essex, E. A., and Moldwin, M. B.: Ionosphere dynamics over the Southern Hemisphere during the 31 March 2001 severe magnetic storm using multi-instrument measurement data, Annales Geophysicae, 23, 707–721, 2005.

Zapfe, B. D., Materassi, M., Mitchell, C. N., and Spalla, P.: Imaging of the equatorial ionospheric anomaly over South America - A simulation study of total electron content, Journal of Atmospheric and Solar-Terrestrial Physics, 68, 1819–1833, 2006.