CHALLENGES IN TOPSIDE IONOSPHERIC MODELLING OVER SOUTH AFRICA

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

DOCTOR OF PHILOSOPHY

---

of

RHODES UNIVERSITY

---

by

Patrick Sibanda

March 2010

Abstract

This thesis creates a basic framework and provides the information necessary to create a more accurate description of the topside ionosphere in terms of the altitude variation of the electron density (Ne) over the South African region. The detailed overview of various topside ionospheric modelling techniques, with specific emphasis on their implications for the efforts to model the South African topside, provides a starting point towards achieving the goals. The novelty of the thesis lies in the investigation of the applicability of three different techniques to model the South African topside ionosphere: (1) The possibility of using Artificial Neural Network (ANN) techniques for empirical modelling of the topside ionosphere based on the available, however irregularly sampled, topside sounder measurements. The goal of this model was to test the ability of ANN techniques to capture the complex relationships between the various ionospheric variables using irregularly distributed measurements. While this technique is promising, the method did not show significant improvement over the International Reference Ionosphere (IRI) model results when compared with the actual measurements. (2) Application of the diffusive equilibrium theory. Although based on sound physics foundations, the method only operates on a generalised level leading to results that are not necessarily unique. Furthermore, the approach relies on many ionospheric variables as inputs which are derived from other models whose accuracy is not verified. (3) Attempts to complement the standard functional techniques, (Chapman, Epstein, Exponential and Parabolic), with Global Positioning System (GPS) and ionosonde measurements in an effort to provide deeper insights into the actual conditions within the ionosphere. The vertical Ne distribution is reconstructed by linking together the different aspects of the constituent ions and their transition height by considering how they influence the shape of the profile. While this approach has not been tested against actual measurements, results show that the method could be potentially useful for topside ionospheric studies. Due to the limitations of each technique reviewed, this thesis observes that the employment of an approach that incorporates both theoretical considerations and empirical aspects has the potential to lead to a more accurate characterisation of the topside ionospheric behaviour, and resulting in improved models in terms of reliability and forecasting ability. The point is made that a topside sounder mission for South Africa would provide the required measured topside ionospheric data and answer the many science questions that this region poses as well as solving a number of the limitations set out in this thesis.

Acknowledgements

This thesis work was done at the Hermanus Magnetic Observatory (HMO) during the years 2007-2009. I therefore wish to thank the HMO for providing the space, computer and internet resources. I also thank everyone at the HMO for being like family to me during this time. I would like to thank my suppervisor Dr. Lee-Anne McKinnell for her guidance throughout this work, for believing in me and encouraging me in the difficult moments. Many thanks go to two people who helped shape this thesis, Elda Saunderson and Auntie Corrina. Thank you for helping me with the grammar and guiding me on this, I will forever be greatful because I have gained knowledge that surely plays a part in defining my future. I also thank my colleagues and fellow students at the HMO, they all played a special part in being such good friends, and for interesting discussions especially during coffee breaks. I was financially supported by the HMO and the National Research Foundation (NRF). Additionally, I received travel support for attending several conferences and workshops from the International Union of Radio Science (URSI), the National Science Foundation (NSF) of the USA and the NRF. Finally, I would like to thank my parents, Gelly and Raymond, for the support and understanding they have given me during these years, without whom this thesis would probably not have been possible. All the glory and honour goes to my God for taking me out of that little village and bringing me this far, leading me every step of the way and working through everyone who has participated in making this work possible.

Preface

Parts of this thesis are published in the publications listed below, and the text in those sections may look identical to the published work. Efforts were made to revise the text and acknowledgement are made to these publications in various sections of the thesis.

- Sibanda, P. and McKinnell, L. A., 'Evaluating the IRI topside model for the South African region: An overview of the modelling techniques,' *Advances in Space Research*, **44**: pp. 707-714, 2009.

- Sibanda, P. and McKinnell, L. A., 'The applicability of existing topside ionospheric models to the South African region,' *South African Journal of Science*, **105**: pp. 387-390, 2009.

Dedication

I dedicate this thesis to Gelly my mum, Raymond my dad and Mrs Sianchongwe my primary school teacher for the special part they played to shape my academic life in a way I can never explain. They established the strong foundation upon which the work of this thesis rests. My parents gave me all the support from my tender years in grade 1 at Mumbwa BOMA primary school and have continued with this support every step of the way. Mrs Sianchongwe played her part and tought me from grade 1 through to grade 7. She tought me how to hold the pencil and use it to write and how to count numbers and use them to solve problems. Mum, Dad and Mrs Sianchongwe, this is the foundation I will cherish all my life.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

The ionosphere is an important part of our solar-terrestrial environment that is highly dynamic and continues to command much interest in many scientific and technological applications. It is the ionised region of the Earth's upper atmosphere where free electrons occur in high enough densities to influence the propagation of radio waves significantly. This thesis is focused on the topside part of the ionosphere above about 350 km, an area that has not yet been fully characterised. The primary science objective is to obtain a better description of the topside ionosphere in terms of the altitude variation of the electron density over the South African region.

In recent years, society has become increasingly reliant on technological systems such as telecommunication techniques and systems that depend on satellite applications. There has also been a tremendous increase in the use of technological applications of Global Navigation Satellite System (GNSS) whose performance depends on the state of the ionosphere. These applications involve the transmission of radio signals from the ground to the satellites which pass through the ionosphere. The ionosphere causes delay in the electromagnetic signals as they propagate through it. The global understanding of the state of the ionosphere in terms of the height distribution of the electron density and the accuracy of critical frequency maps is therefore a key point in the development of satellite communication applications and also navigation applications that are based on GNSS. Long-term models of the state of the ionosphere are essential in communication operations and radio propagation for calculating communications routes, and planning HF radio propagation circuits (Bradley and Dick, 1997). Ionospheric modelling is also important in providing us with a deeper understanding of the characteristics of the ionosphere and the phenomena that take place in it which is critical in assessing the effects of the ionosphere on Earth-space links. Ionospheric models are also important in the use of the ionosphere as a natural communication channel in long-haul communications that are based on ionospheric reflections.

Ionospheric modelling in South Africa has concentrated on the bottomside ionosphere with the use of the data from the three South African ionosonde stations. A national

model, the South African Bottomside Ionospheric Model (SABIM) was developed using the data from these stations. The model predicts the bottomside electron density profile up to the height of the F2-peak, h_mF2, given the geographical location, time of day and season, solar activity and geomagnetic activity (McKinnell, 2002). There is a need to expand this model to include the topside ionosphere above the F2-peak. This will make the model more representative of the complete ionosphere over the South African region and also provide it with an added value in being able to predict the Total Electron Content (TEC) value to around 2000 km.

Many scientists around the world have studied the topside ionosphere and attempted to model it, (e.g Bilitza *et al.* (2006); Depuev and Pulinets (2004); Huang and Reinisch (2001); Kutiev *et al.* (2006); Pulinets *et al.* (2002); Radicella and Zhang (1995); Reinisch and Huang (2001); Stankov *et al.* (2002)), however, none has concentrated on the South African topside ionosphere. Currently very little is known about the ability to model the topside over the South African region and many global ionospheric models have always been insufficient for the South African region due to a paucity of available measurements (McKinnell and Poole, 2004).

In the recent years, a wide range of topside ionospheric models that predict the variation of the electron density with height under different conditions have been developed. Schunk and Sojka (1992) presented an overview of various approaches to ionospheric modelling which include an empirical approach, analytical approach, theoretical approach and approaches involving data assimilating techniques. In a similar study, Cander *et al.* (1998) reviewed the progress in ionospheric modelling and described the various types of modelling approaches with a view to assessing their applicability to space weather forecasting. Bilitza *et al.* (2006) conducted a comprehensive review of the different techniques of topside modelling, and how they have been applied in various existing models, with an emphasis on their implication for the International Reference Ionosphere (IRI) model. They described the major modelling efforts focusing on how these efforts can help in improving the representation of the electron density profile prediction in the IRI model. In earlier studies, Bilitza (2004) evaluated the IRI topside model. A comparison of the model results with Alouette and International Satellites for Ionospheric Studies (ISIS) topside sounder data showed that the model systematically overestimates the measurements. Part of this study follows on these reviews to provide an overview of topside modelling with a specific focus and emphasis on the implications of these various modelling techniques for the topside modelling efforts in South Africa.

1.1 Objectives

This thesis is devoted to studying the topside ionospheric electron density focusing on the abilities to model the topside ionosphere over the South African region and exploring ways of improving the characterisation of the altitude distribution of the electron density (Ne) in this region. The main objective of this thesis is to provide a comprehensive understanding of topside ionosphere modelling in South Africa to serve as a starting point in developing a topside ionosphere model that is suitable for this region. This was motivated by the fact that during the last few years the need for increased reliability of technological systems, whose performance depends on the state of the ionosphere, has significantly increased in South Africa, but the ionospheric studies have concentratred only on the bottomside which is currently well represented in SABIM. Not much has been done about the topside which also plays a significant role in the effects the ionosphere has on the performance of technological systems, such as the GNSS signals, and therefore cannot be ignored. The long-term goal is to develop a topside ionosphere model that best represents the electron density variation in the topside ionosphere over the South African region and smoothly connect such a model to the bottomside profile predicted by SABIM. It is expected that this thesis will provide the information needed to decide on a program for this work in the future.

1.2 Tasks undertaken

As a first step towards achieving the objectives, a comprehensive review and study of the current status of topside modelling and the techniques used in the various topside modelling efforts was undertaken in order to explore the possibility of their application in modelling the topside ionosphere over the South African region. In the same work, an evaluation of how well the IRI model represents the topside ionosphere over the South African region under different conditions was presented and the results of the study are published in Sibanda and McKinnell (2009a). The evaluation involved the IRI-2001 version of the IRI which was the latest version as at the stage the study was undertaken. The results of the IRI topside electron density prediction were compared with the measured profiles from the ISIS-2 topside sounder. The study found that the IRI-2001 model significantly overestimated the measurements at higher altitudes (above about 800 km). It was further observed that the model require updating with the measured F2-peak parameters, $foF2$ and h_mF2, otherwise the deviation with the actual measurements is very large.

A detailed study of the available measured data sets that can be useful for topside

modelling in the South African region was undertaken. This revealed that the coverage of the South African region is very poor in the measurements and this poses the greatest challenge regarding modelling the topside ionosphere in this region.

1.3 Thesis outline

This chapter has introduced the subject and objectives of the thesis. Chapter two presents the background knowledge on the ionosphere in general and the topside ionosphere in particular. An overview of the various approaches and techniques used in modelling the topside ionosphere and how they are applied in various existing topside ionosphere models is presented. A great deal of analysis of the different topside modelling techniques and approaches, with the prime goal of identifying the approach that can best be applied for topside modelling in the South African region where measurements are scarce, was performed, and these are given in separate chapters. Chapter three deals with the empirical approach based on the available topside sounder data sets using an Artificial Neural Network approach in an effort to determine the applicability of this approach in modelling the topside over the South African region with the available data. Coverage of the datasets is presented and it is noted that coverage of the South African region in the database is very limited and this presented the greatest challenge to the application of this approach for the South African situation. Chapter four presents the theoretical approach option. The principle of diffusive equilibrium in the ionosphere and how this principle can be used to model the distribution of the charged particles in the topside ionosphere is discussed in detail. A number of models which are discussed in the literature apply this approach, but they were not available for testing during the preparation of this thesis. Attempts are made, following this approach, to create a representation of the topside ionosphere over South Africa. The results from these efforts are presented as well as the difficulties and limitations encountered. Chapter five deals with the empirical modelling approach involving data assimilation techniques and presents the efforts made to develop a topside model following the approach first presented by Stankov and Muhtarov (2001); Stankov *et al.* (2003) and Stankov *et al.* (2002) which is based on the use of Global Positioning System (GPS) derived Total Electron Content (TEC), ionosonde measurements and an empirically obtained upper transition (UTH). The approach presents a potential for improving ionospheric characterisation in the South African region with the availability of GPS data from the wide network of GPS receivers installed across the country and the neighbouring countries. However, there are still many challenges and limitations for this approach and these are discussed in detail in the chapter. Finally in chapter six, a

summary of the thesis and a discussion of the major observations and conclusions are presented. Future directions are also discussed and recommendations are given for what should be taken into account in any subsequent topside ionosphere modelling work in order to further this work and achieve the ultimate goal, to develop a topside ionosphere model for this region.

Chapter 2

Theory and Background

This chapter presents the basic principles and theoretical background of the ionosphere. Section 2.1 discusses the general structure and characteristics of the ionosphere and its variabilities. The rest of the chapter focuses, in more detail, on the characteristics of the topside ionosphere. Various methods for measuring the topside ionosphere and the importance of its characterisation are discussed. This leads to the overview of various approaches to modelling the topside ionosphere and the techniques often used, which in general, involve either an empirical approach or a theoretical approach. The review of the modelling techniques is done with a specific focus and interest on the implications of these techniques for topside modelling efforts in South Africa, in line with the desire to expand existing bottomside ionospheric models to include the topside region, and provide a more complete representation of the ionosphere over this region.

2.1 The ionosphere and its characteristics

The ionosphere is a part of the Earth's space weather system that is generally accepted to begin at about 60 km above the Earth's surface. The upper boundary is not clearly defined however, it is generally accepted to range up to a few thousand kilometers at the start of the plasmasphere where H^+ becomes the dominant ion species (Bassiri and Hajj, 1992; Verronen, 2006). It is characterised by free electrons and positively charged ions produced by the interaction of solar radiation with the atmospheric constituents. The ionisation (the breaking away of the electrons from the atoms and molecules) originates from the absorption of solar radiation (mainly ultraviolet, extreme ultraviolet and X-rays) by neutral gases (McNamara, 1991), the major one being atomic oxygen (O, ionised to O^+) at about 60 - 300 km altitude (Havens *et al.*, 1954; Verronen, 2006). Other gases available for ionising include hydrogen (H, to H^+) and nitrogen monoxide (NO, to NO^+) (Verronen, 2006). With decreasing altitude towards the surface of the Earth, the solar radiation encounters an increasing number of the neutral atmospheric molecules, and absorption increases while the intensity of the radiation decreases. This leads to a layer of maximum ionisation occurring typically at altitudes of between 200 and 400 km (García-Fernández, 2004).

The electron density is an ionospheric variable of great importance in the characterisation of the structure of the ionosphere and it varies substantially with height. Thus the ionosphere is categorised into three broad regions: the bottomside ionosphere, the topside ionosphere and the plasmasphere. These regions may be further divided into several regularly occurring layers. Figure 2.1 is a schematic illustration of the various regions and the layers of the ionosphere.

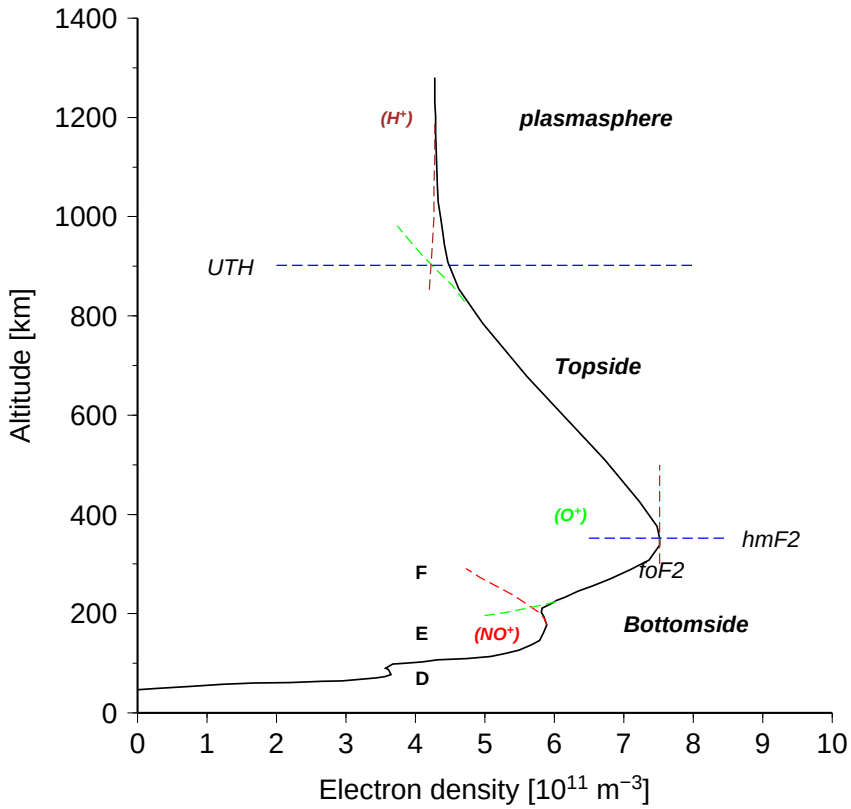


Fig. 2.1: Schematic illustration of the ionosphere showing the various regions (bottomside, topside and plasmasphere) of the ionosphere and the ion species associated with the formation of the different ionospheric regions

In the bottomside ionosphere, ionisation appears at different ionospheric levels, producing a series of distinct stratified layers called D, E and F layers (McNamara, 1991), which may be identified by their interaction with radio waves and by the distinct source of ionisation that causes them. Chemical reactions, such as photoionisation, ion-molecular reactions, and electron-ion recombination are responsible for the plasma distributions in the bottomside ionosphere.

- The F-layer is ionised by ultraviolet and extreme ultraviolet radiation and is the predominant layer of the ionosphere which contains the bulk of the electron content (Havens *et al.*, 1954). This layer begins at about 150 km above the Earth's surface

to approximately 400 km (Rishbeth and Garriott, 1969). Maximum ionisation in the ionosphere occurs in the F-layer at the maximum height of the F2-layer and marks the peak of the bottomside ionosphere. During the daytime, the F-layer is split up into the F1 and the F2-layers which are produced by extreme ultraviolet radiation (McNamara, 1991). At night the two F-layers combine into the F region. The F-layer affects HF radio waves with higher frequencies that can penetrate the lower layers of the ionosphere.

- Below the F region is the E-layer produced by soft X-rays. It ranges from about 90 km to 130 km (McNamara, 1991), and is present throughout the day with a much reduced electron density at night.

- The region closest to the Earth's surface that ranges from about 60 km to about 90 km is the D-layer. It is produced by hard X-rays from the Sun during the daytime and quickly disappears at night.

The bottomside ionosphere can be measured quite accurately with ground-based ionosondes which record data as ionograms and is well described in various global empirical models such as the International Reference Ionosphere (IRI) model (Bilitza, 1990), and regional models such as the South African Bottomside Ionospheric Model (SABIM) (McKinnell, 2002).

Above the F2-layer peak, lies the topside ionosphere (discussed in more detail in section 2.2). The upper boundary of the topside ionosphere is not well defined since the electron density thins into the plasmasphere and subsequently into the planetary plasma. It is generally assumed to end at around 1500 km where O^+ becomes less dominant and H^+ becomes more dominant. The region above the topside ionosphere where the majority of the ions correspond to H^+, is known as the protonosphere or "plasmasphere".

2.1.1 Variabilities of the ionosphere

The Sun and its activities are the main driver of ionisation in the ionosphere (Carpenter, 2004). Thus, the ionospheric structure and the electron density distribution varies considerably with time (diurnally, seasonally), space (geographic location), Sun's activity (sunspot cycle) and with other solar-related ionospheric and geomagnetic disturbances.

2.1.1.1 Diurnal variation

The ionosphere exhibits a diurnal variation where ionisation is normally higher during the day and lower at night. As the Earth rotates with respect to the Sun, the intensity of

the solar radiation increases and the ionisation also increases to a maximum at around noon. During the afternoon, ionisation begins falling due to electron loss and with darkness the ionosphere becomes higher and thinner. Through the night, maximum frequencies gradually decrease, reaching their minimum just before dawn.

2.1.1.2 Seasonal variation

As the Earth revolves around the Sun, a seasonal cycle is generated, determined by which hemisphere the Sun is overhead. It is summer on the hemisphere where the sun is overhead and winter on the other hemisphere. The Sun is overhead at the equator around the time of the equinoxes. This seasonal cycle causes a corresponding seasonal and spatial variation in the global ionospheric structure. The D-, E- and F1-region electron densities are greater in summer than in winter (McNamara, 1991). However, the variation in F2-region electron densities is more complicated. In both hemispheres, electron densities generally peak around the equinoxes (March and September) (García-Fernández, 2004). Around solar minimum the summer noon electron densities are, as expected, generally greater than those in winter, but around solar maximum winter frequencies tend to be higher than those in summer a phenomenon called seasonal anomaly. In addition, frequencies around the equinoxes (March and September) are higher than those in summer or winter for both solar maximum and minimum.

2.1.1.3 Solar cycle variation

Measurements of the sunspot number, an indicator of solar activity, show that the Sun goes through a periodic rise and fall in activity with a period of about 11 years. Since ionisation is driven by solar electromagnetic and corpuscular radiation, it is a function of solar activity. At solar maximum the Sun emits more radiation and therefore ionisation is greater. Thus the critical frequencies of the ionospheric layers are greater during high solar activity periods.

2.1.1.4 Latitudinal variation

The intensity of radiation and the daily production of free electrons decreases with increasing latitude, since with increasing latitude, the solar radiation strikes the atmosphere more obliquely. With respect to geographical location, the ionosphere may be characterised broadly into three regions (low-latitude, mid-latitude and high latitude) that have different properties.

In the *low-latitude region*, the general form of the ionosphere is distorted with a depletion of electron density over the geomagnetic equator and large enhancements occurring around $\pm 20^{\circ}$ north or south of the magnetic equator. This phenomenon is caused by

the "fountain effect": an electrodynamic lifting of the plasma which drifts upwards until the pressure and gravity forces are high enough and the plasma is then forced back through the magnetic field lines to higher latitudes. This effect is a consequence of the fact that the geomagnetic field runs horizontally over the geomagnetic equator causing strong electromagnetic forces that lead to an abnormally large electrical conductivity over the equator and a strong electric current (an "electrojet") that flows in the E and the F regions.

The *mid-latitude ionosphere* shows the least variations. It is considered to be the quietest but it is known that some seasonal anomalies and some short term variations such as the main ionospheric trough (MIT) driven by solar and geomagnetic storm inputs in the auroral ionosphere occur in these regions (Yizengaw, 2004). Also, the interaction of equatorward propagating disturbances with the ionospheric plasma and also with the constituents of the neutral atmosphere affects the movement of ions, and the balance between production and loss in these regions (García-Fernández, 2004). These disturbances affect radio wave propagation. The mid-latitude ionosphere has been explored the most because most ionosphere observing instruments are located in these regions and it is the best understood region. Ionisation is produced almost entirely by ultra-violet and X-ray emissions from the Sun, and is depleted again by chemical recombination processes.

In the *high latitude regions* the electron densities are considerably lower than in the lower latitudes since the solar radiation strikes the atmosphere in this region at a more oblique angle. The high latitude ionosphere is rich in plasma instabilities because of its susceptibility to external influences caused by the fact that the geomagnetic field runs nearly vertical, and the magnetic field lines connect to the outer part of the magnetosphere which is driven by the solar wind. Energetic particle emissions from the Sun can precipitate into the ionosphere and produce additional ionisation leading to sporadic events adding to the complexity of the ionosphere in this region (García-Fernández, 2004).

2.2 The topside ionosphere

The topside ionosphere extends from the peak of the F2-layer and thins out into the plasmasphere at an altitude of about 1500 km where H^+ becomes the dominant ion and the ionic reactions have little practical importance (Banks *et al.*, 1976). Strictly speaking, the upper boundary of the topside ionosphere is not distinctly defined. In this study, like in other studies (e.g Yoshimura *et al.* (2005) and Marinov *et al.* (2004)), it is

accepted to be at the Upper Transition Height (UTH), the height at which the densities of the O^+ and H^+ ions are equal. This definition assumes O^+ and H^+ ions are the major ion species which play a significant role in the topside ionosphere. He^+ is assumed to be the minor ion. However, Heelis *et al.* (1990) found theoretical evidence that sometimes the He^+ density rises above the H^+ density in this region, an effect that is referred to as winter nighttime He^+ bulge. Both He^+ and H^+ ions are strongly coupled with their corresponding neutral densities but the physical processes controlling their densities are quite different. The focus in this study will be on the topside ionosphere, with O^+ and H^+ as the dominant ions.

Various photochemical processes are responsible for the production and destruction of the electron-ion plasma in the ionosphere, as pointed out above. In the topside ionosphere the density of the neutral molecules and atoms is much less compared to their densities in the lower ionosphere. Therefore, most of the production and loss of the ionisation occurs in the lower ionosphere, especially around the F-region (Webb and Essex, 2000). The electron density distribution in the upper ionosphere is therefore closely linked to the dynamics of the lower ionosphere. Due to the low density of the neutral atoms in the topside ionosphere, diffusion is rapid enough to cause the plasma to diffuse away from the F-region ionosphere a significant vertical distance in the order of an atmospheric scale height (about 50 km) within its life time of a few hours (Rishbeth, 1975). Thus, the dynamics of the topside ionosphere in terms of the distribution and the relative amounts of the O^+ and H^+ densities are influenced by plasma transport processes, field-aligned plasma flows (Venkatraman, 1999) and chemical processes (Rishbeth and Borron, 1960). The F-region is both a source and a drain of ionisation for the topside ionosphere.

2.2.1 Measuring the ionosphere

Early studies of the ionosphere were limited to the bottomside ionosphere, based on observations of the electron density using ground-based ionosondes. Ionosondes have provided much of the knowledge known about the ionosphere, especially the bottomside ionosphere. They have been used since the 1920s to study the spatial and temporal characteristics of the ionospheric electron density (Havens *et al.*, 1954). An ionosonde is a pulsed radar system that sweeps in frequencies from around 1 to 30 MHz (Hunsucker and Hargreaves, 2003) approximately every 32 seconds vertically into the ionosphere and the signals are reflected off the ionospheric layers. The lowest frequency at which the signal passes through the ionosphere is called the critical frequency. By measuring the time taken for a reflection of the signals from the ionosphere, the electron density, as a

function of altitude above the ionosonde, can be deduced.

Due to the properties of high frequency (HF) radio wave reflections the topside ionosphere can not be probed by ground-based ionosondes, which are limited to sounding the bottomside ionosphere. Any radio signal that has a frequency greater than the critical frequency passes beyond the peak height of the F2-region and is not reflected, and thus provides no information. The advent of satellite mounted ionosondes (topside sounders) in the 1960s added a new dimension to the study of the ionosphere. Topside sounders perform measurements of the topside ionosphere comparable to the measurements of the bottomside ionosphere performed by the ground-based ionosondes. Other techniques that permit the measurement of the region above the level of the F2-layer peak include the use of rockets and Incoherent Scatter Radars (ISR) systems. Instruments can be mounted on rockets to make direct measurements of the ionospheric parameters, such as relative concentration of the various ion species and electron/ion temperatures. However, such measurements are brief and can only be made along the trajectory of the rocket. Ionospheric observations using rockets represent "direct" methods where the assumptions and approximations that must be made involve only local parameters, and therefore only localised measurements are recorded that can lead to localised conclusions. Incoherent Scatter Radars provide another method for determining electron density distribution in the topside ionosphere. In this method, energy is propagated into the ionosphere from an extremely powerful pulsed radar. Some of the echos received are due to the scattering of waves from the individual electrons in the ionosphere. The power of the returned signal is directly proportional to the electron density.

In the case of a topside sounder, as the satellite traverses its orbit, the topside sounder records the ionogram (a picture of reflections and resonances of the swept frequencies against apparent range) traces similar to those recorded by bottomside ionosondes. From the ionograms, the electron density at a particular time and place within the height range between the satellite height and the peak of the F2-layer can be deduced by use of an inversion procedure. The inversion of topside ionograms is less complicated than that of bottomside ionograms which have to take into account the valley region between the E- and F-layers. In the case of the topside ionograms, the inversion procedure uses the electron density at the starting point, which is usually known from simultaneous measurements by an on-board in-situ probe (Bilitza, 1994).

2.2.1.1 Topside sounder missions

Comprehensive observations of the topside ionosphere began in the early 1960s when the topside sounder satellites Alouette and ISIS where launched and provided global observations of the topside ionosphere. The main goal of the satellites was to perform topside soundings of the ionosphere to determine electron density profiles from the F2-layer peak to the altitude of the satellite. The joint USA and Canadian project designed and operated the first generation of topside sounder satellites; the first one was Alouette-1 which was launched and became operational in 1962 (Hartz, 1964). The second was the Alouette-2 satellite which was launched in 1965 (Florida, 1969) as a backup in case Alouette 1 malfunctioned. The major limitation with this generation of satellites was that they had no on-board storage capacity, and relied on the network of ground telemetry stations to collect and transmit data. Data could only be sent down when the spacecraft was in the line of sight of a telemetry station (Bilitza, 2001). The Alouette program was successful and that led to the development of the International Satellites for Ionospheric Studies (ISIS) series, which had data storage capacity on-board. This program was also a joint venture between the USA and Canada. On this program, the ISIS-1 spacecraft was launched into an elliptical orbit with a perigee of 550 km and an apogee of 3500 km in 1969 (Gillies, 2006). Later, in 1971, ISIS-2 was launched into a nearly circular orbit of 1400 km.

Other known topside sounding missions include:

- The Japanese operated satellites ISS-a & -b which were launched in 1976 and 1978 respectively (Benkova, 1990). ISS-a operated only for one month, at which stage its power systems failed. On the other hand, ISS-b did not carry enough on-board memory, and therefore did not collect much data (Wakai and Matuura, 1980).

- Intercosmos 19 (denoted IK-19) satellite designed and operated by the Russian Institute of Terrestrial Magnetism, Ionosphere and Radiowave propagation (IZMIRAN), was launched in 1979 and operated until 1982 (Benkova, 1990). Later in 1986 the Russian Academy of Sciences laboratories launched the Cosmos-1809 satellite (Shuiskaya *et al.*, 1990) carrying a payload similar to that which was carried by the IK-19. It was placed into a 960 km circular orbit and operated until May 23, 1993.

The topside sounder measurements from these satellites provided a valuable data source for studies of the global morphology of the topside ionosphere, although only a small fraction of the ionograms were processed to electron density profiles (Bilitza, 1994).

Empirical modelling of the topside electron density has relied on data from these missions. However, current knowledge of the structure of the topside ionosphere is not sufficient yet since coverage of relevant geophysical conditions in the processed database is limited (Reinisch and Huang, 2001).

Table 2.1 shows a list of known topside sounder missions, their launch dates and the period for which the data is available.

Table 2.1: Topside sounder satellites (Bilitza, 1994)

Satelite name	Operated by	Launch date	data available until	Orbit altitude /km
Alouette-1	USA/Canada	Aug 1962	Sep 1972	1000
Alouette-2	USA/Canada	Nov 1965	Jan 1975	500 - 3000
ISIS-1	USA/Canada	Jan 1969	Dec 1981	2000 - 3500
ISIS-2	USA/Canada	Apr 1971	Jan 1983	1400
ISS-b	Japan	Feb 1978	Apr 1983	1100
IK-19	Russia	Feb 1979	Aug 1981	500 - 1000
Cosmos-1809	Russia	Dec 1986		960

2.2.1.2 Characterising the topside ionosphere

Reliable information on the spatial distribution of the electron density in the ionosphere is essential for various ionospheric scientific studies as well as for associated practical applications, such as the estimation and correction of Global Navigation Satellite System (GNSS) propagation delays and investigation of space-weather effects on telecommunication systems, amongst others. Considerable efforts have therefore been concentrated on modelling the ionspheric electron density. Maps of ionospheric electron density, if available in real-time, provide useful input to space-weather forecasting.

In recent years, progress in the characterisation of the structure and dynamics of the topside ionosphere has been rapid and many studies (discussed in the following section) have resulted in a wide range of models that provide a global representation of the altitude distribution of the electron density in the topside ionosphere. Due to the complicated nature of the topside ionosphere, there have been numerous approaches to modelling it. The modelling efforts can be placed into two categories: (1) Empirical models based on available world-wide ionospheric measurements, (2) Theoretical or first principle models based on the physics and chemistry that drive the ion distribution in the ionosphere. The section that follows describes these modelling approaches and also reviews the current status of topside modelling.

2.2.2 Modelling the topside: Empirical approach

Over the years, empirical models have been used widely to describe ionospheric structure and behaviour. Typically, with empirical modelling, simple mathematical functions are fitted to measured data which are usually binned with appropriate indices. They attempt to extract systematic ionospheric variation from past data records, hence, empirical models represent average conditions and not instantaneous ionospheric conditions (Schunk and Sojka, 1992). Empirical modelling of the topside ionosphere has been especially challenging due to the limited coverage of the relevant geophysical conditions in the available measured data sets. Since the end of the topside sounder missions described above, there has not been any new and recent measurements.

There have been considerable efforts by various groups (Bilitza *et al.*, 2006; Depuev and Pulinets, 2004; Huang and Reinisch, 2001; Kutiev *et al.*, 2006; Pulinets *et al.*, 2002; Radicella and Zhang, 1995; Reinisch and Huang, 2001; Stankov *et al.*, 2002) to develop empirical models of the spatial distribution of the electron density (*Ne*) especially the height profile (*Ne(h)*) in the topside ionosphere. The modelling efforts involved the application of different techniques. Several mathematical functions such as Chapman, Exponential, Sech-Squared (Epstein) and Parabolic functions have been used to reconstruct the altitude profile of the electron density in the topside ionosphere (Bent *et al.*, 1972; Booker, 1977; Chapman, 1931; Di Giovani and Radicella, 1990; Huang *et al.*, 2002). A description of each of these functions follows:

(a) The Chapman function

The general form of the Chapman layer is:

$$N(h) = N(h_m)\left(e^{c[1-z-e^{-z}]}\right) \tag{2.1}$$

where $z = \frac{h-h_m}{H}$. This function can take one of two forms: an α-Chapman where $c = 0.5$ or a β-Chapman where $c = 1$, depending on assumptions related to the electron recombination theory (Booker, 1956; Hargreaves, 1992). In the β-Chapman form, the assumption is made that the electron loss is through the interaction with the neutral particles while the α-Chapman formulation assumes that the electrons recombine directly with the positive ions and that there are no negative ions (Davies, 1996; Stankov *et al.*, 2003).

(b) Sech-squared function

The sech-squared function, also known as the Epstein function is defined as:

$$N(h) = N(h_m) sech^2 \left(\frac{h - h_m}{2H} \right) \tag{2.2}$$

(c) Exponential function

The exponential function is defined as:

$$N(h) = N(h_m) exp \left(-\frac{h - h_m}{H} \right) \tag{2.3}$$

(d) Parabolic function

The parabolic function is defined as:

$$N(h) = N(h_m) \left(1 - \left[\frac{h - h_m}{2H} \right]^2 \right) \tag{2.4}$$

In each function, h_m is the peak density height and H is the ionospheric scale height.

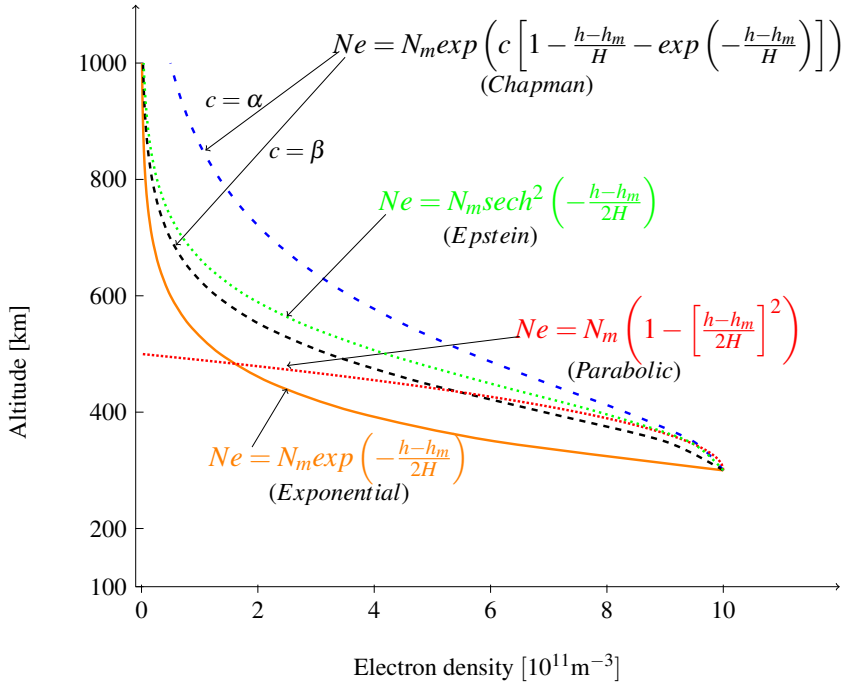


Fig. 2.2: The shape of the topside electron density profile computed with different profile functions

Figure 2.2 presents a comparison between vertical electron density profiles obtained with the different mathematical functions given above for the same ionospheric scale height of 100 km. This shows how the shape of the electron density profile is dependent on the chosen topside profiler. Using the same ionospheric scale height (H), the shape of the topside electron density profile differs considerably.

This demonstrates that the accuracy of reconstructing the topside electron density profile depends on the topside profiler model applied and how the ionospheric scale height is determined. Stankov *et al.* (2003) evaluated the accuracy of these functions in reproducing measured data. Using experimental data from the Atmospheric Explorer-C (AE-C) satellite they found that the exponential and the α-Chapman functions perform well for the daytime measurements, while the Epstein function was better at representing the nighttime measurements.

Generally, many topside ionosphere modelling efforts involved using a unique approach to determine the ionospheric scale height and then using one of the standard profile functions presented above to reconstruct the electron density profile.

2.2.2.1 Ionospheric scale height

The ionospheric scale height is the key and inherent ionospheric parameter for each profile function (Liu *et al.*, 2006; Stankov *et al.*, 2003). It provides a measure of the shape of the altitude dependence of electron density (Liu *et al.*, 2007b) and intrinsically connects to the ionospheric dynamics such as plasma temperatures. There are various definitions of ionospheric scale height in published literature (e.g. Kutiev *et al.* (2006) and Huang and Reinisch (1996)). Liu *et al.* (2007a) undertook a comprehensive study of the ionospheric scale height that differentiated among the various forms of scale heights specifically identifying three forms: (i) plasma scale height, (ii) Vertical Scale Height (VSH) (Kutiev *et al.*, 2006) and (iii) effective scale height (Huang and Reinisch, 1996). They adopted the following definitions of scale heights:

(i) The plasma scale height (H_p) is defined in terms of its intrinsic connection to ionospheric dynamics such as plasma temperature (T_i and T_e respectively the ion and electron temperatures) and plasma compositions (Luan *et al.*, 2006). It is given as

$$H_p = k_b(T_i + T_e)/m_i g \tag{2.5}$$

where k_b is the Boltzmann constant, g the acceleration due to gravity and m_i the ion mass.

(ii) The vertical scale height (VSH) is related to the gradient of the measured electron density profiles. It is the vertical distance in which the topside electron density changes by a factor of $e(\approx 2.718282)$. This can be expressed mathematically as

$$VSH = -Ne\frac{dh}{d(ln(Ne))} \tag{2.6}$$

if an exponentially decaying electron density profile is assumed. Where Ne is the electron density and h is the height. The VSH can be obtained by fitting equation 2.6 to the measured Ne profiles.

(iii) The effective scale height (H_m) is the scale height used in fitting the electron density profiles with the α-Chapman function (e.g. Reinisch *et al.* (2004)).

Most empirical topside modelling efforts have relied more on the VSH or the effective scale height (H_m) than the plasma scale height (H_p) to provide a measure of the altitudinal dependence of the ionospheric electron densities (Liu *et al.*, 2007a). Thus, many of the topside ionosphere empirical modelling studies have been centred around calculating the VSH, which is then used in the reconstruction of the electron density profiles. Several data sources have been used including GPS-TEC and radio occultation measurements, ground-based ionosondes, topside sounders, and incoherent scatter radars (ISR). The methods have included:

(i) Calculating VSH from GPS derived TEC (Jakowski *et al.*, 2002; Stankov *et al.*, 2003) and then using this to reconstruct the electron density profiles up to the satellite height.

(ii) Obtaining VSH from ground-based ionosonde measurements (Huang and Reinisch, 2001). An estimate of the topside scale height is derived from the shape of the bottomside profile at the F2-peak. This is then used to reconstruct the topside profile. In a recent study, Reinisch *et al.* (2007), improved this earlier approach to use a scale height that varies with altitude. The scale height varies slowly with altitude around the F2-peak and then increases rapidly at the UTH. They used a hyperbolic tangent given by equation 2.8 to represent this behaviour, and discussed in more detail in section 2.2.2.2.

(iii) Calculating the VSH from topside sounder data. Depuev and Pulinets (2004) used data from the Intercosmos-19 satellite. Their approach showed the importance of the longitudinal variation of the scale height for global modelling. Kutiev *et al.* (2006) used the ISIS-2 and the Intercosmos-19 topside sounder data to extract the

vertical O^+ scale height which they defined as the lowest gradient of the measured electron density profiles.

The sections that follow highlight some of the topside modelling efforts with particular emphasis on the techniques applied. Some of the models that were available are compared to measured data to determine how they represent the topside ionosphere over the South African region.

2.2.2.2 Empirical models

This section discusses some available empirical models of the topside ionospheric electron density focusing on the techniques used to create the models. The purpose is to identify the most adequate approach to use in the South African efforts to develop a topside ionosphere model suitable for the South African region.

The International Reference Ionosphere (IRI) model

The IRI is an international project sponsored by the Committee on Space Research (COSPAR) and the International Union on Radio Science (URSI) (Bilitza, 1991b). The IRI model is an empirical standard model of the ionosphere that describes the average electron density, ionospheric electron, ion and neutral temperatures and ion composition as functions of height, location, local time and sunspot number for magnetically quiet conditions (Bilitza, 1990). It is based on data from the world-wide network of ionosonde stations, incoherent scatter radar, Alouette topside sounders and in-situ measurements made by several satellites and rockets. The IRI model has been updated and improved regularly since it was first created in the 1960s. Annually, special IRI workshops are organised during which the IRI working group meets and discusses improvements to the model.

The representation of the topside electron density profile in the IRI model is based on the ISIS and Alouette topside sounder data sets. The formulation of altitude variation is described by the Booker function (Booker, 1977; Coïsson *et al.*, 2006) where the topside is divided into three altitude segments (Bilitza, 2004; Bilitza *et al.*, 2006) with each of the segments modelled with a constant scale height computed with the Bent *et al.* (1972) model. Rawer *et al.* (1978) used Bent's parameters which were provided in the form of tables and graphs to develop an analytic description of the information contained in these parameters.

When the IRI project started, the topside electron density formulation relied entirely on the Bent *et al.* (1972) model's data compilation owing to the absence of any other directly available data sources. The topside representation in the Bent *et al.* (1972) model was

based primarily on Alouette-1 topside sounder measurements and the parameters in this model were provided in the form of tables and graphs. The Alouette satellite did not have on-board storage capabilites, and the global coverage depended on the distribution of receiving ground stations most of which were located in the northern mid-latitude regions. Furthermore, the Alouette satellite only covered low and medium solar activity periods. As a result the Bent *et al.* (1972) model and the early version of the IRI model showed shortcomings at high solar activities and at low and high latitudes.

A number of studies have evaluated the IRI topside model. For example, Bilitza and Williamson (2000) compared the electron density profiles predicted with the IRI model with the measured electron density profiles from the ISIS satellite. The study reviewed shortcomings in the IRI model. It was found to overestimate the observations mostly in the upper part of the topside ionosphere. Ezquer *et al.* (1998) used simultaneous TEC and F-peak measurements at a station close to the southern crest of the equatorial anomaly and found that the profile shape of the topside in the IRI model was responsible for the discrepancy between the measured data and the model results.

Several recent topside ionospheric modelling studies devoted their efforts to improving the topside representation in the IRI model using different approaches. Some of the studies such as Depuev and Pulinets (2004), used data from topside sounders other than the Alouette and the ISIS sounders. Triskova *et al.* (2002) used data from in-situ measurements. Other studies (for example Kutiev and Marinov (2007); Kutiev *et al.* (1994, 2006); Marinov *et al.* (2004)), considered certain key parameters of the ionosphere such as the scale and transition heights, and modelled these, based on all available topside sounder data. Others (Reinisch *et al.*, 2007) aim at extending the modelling into the plasmasphere focusing on the connection between the topside and the plasmasphere.

Bilitza *et al.* (2006) presented a comprehensive review of the different approaches to topside ionospheric modelling and their implications for improving the topside representation in the IRI model. In an earlier study, Bilitza (2004) used the Alouette and ISIS topside profiles to evaluate the IRI model. Averages of the ratios of measured data versus model results were computed, and based on these ratios, correction factors for the IRI topside model were established. The factors vary with altitude, modified dip latitude and local time. The latest version of the IRI model (IRI-2007) incorporates this corrected topside model. In addition, IRI-2007 also includes the NeQuick model topside electron density formulation as an alternative to using the correction factor option. Both these formulations have used ISIS-2 topside sounder data to improve the representation of the topside ionosphere.

The NeQuick topside model

The topside component of the NeQuick model (Radicella and Zhang, 1995), describes the topside electron density profile using an Epstein layer (Coïsson *et al.*, 2002) with varying scale height given by equation 2.7 (Radicella and Leitinger, 2001):

$$H(h) = \frac{kB_{bot}}{\nu}\left\{\frac{1+12.5(h-h_m)}{100kB_{bot}+0.125[h-h_m])}\right\} \tag{2.7}$$

where B_{bot} is the bottomside thickness parameter, and k is a correction factor that depends on the F-peak density. It applies the DGR "profiler" concept, a concept that uses an Epstein function whose shape is controlled by an empirical parameter connected to the bottomside thickness parameter. The DGR "profiler" concept was first developed by Di Giovani and Radicella (1990) and later modified by Radicella and Zhang (1995). The model is based on the combination of bottomside ionosonde measurements and the ISIS and Intercosmos-19 topside sounder measurements. It assumes a close correlation between topside and bottomside thickness parameters (Bilitza *et al.*, 2006); thus, the scale height (equation 2.7) is given as a function of the bottomside thickness parameter.

The F2 topside and IMAGE/RPI plasmaspheric model

Reinisch *et al.* (2007) developed a new model which is a merge of the ionosonde topside model with the IMAGE/RPI plasmaspheric model. The model uses a varying α-Chapman function (vary-Chap) that takes into account the variation of the scale height with altitude. Reinisch and Huang (2001), Huang and Reinisch (2001) and Reinisch *et al.* (2004) successfully showed how to approximate the topside electron density profile assuming an α-Chapman profile with a constant scale height deduced from ground-based ionosonde measurements around the F2-peak. This procedure is currently used in ionosonde software, such that the complete electron density profile in ionosonde measurements consists of a measured bottomside and a modelled topside component.

In the new and improved approach, the scale height is allowed to vary slowly with altitude around the F2-peak and then increases rapidly at the UTH. They used a hyperbolic tangent to exhibit this behaviour defined as:

$$H(h) = H_T + \frac{H_m - H_T}{tanh(\beta)} tanh\left(\beta\frac{h-h_T}{h_m-h_T}\right) \tag{2.8}$$

where H_T is the scale height at the O^+-H^+ transition (h_T) and β is the shape factor that controls the steepness of the O^+-H^+ transition. This approach makes it possible to merge the F2 topside model with the IMAGE/RPI plasmaspheric model (Reinisch,

2004). The model uses the plasmasphere density profile data measured with the RPI instrument on the IMAGE satellite, and topside sounder data from the ISIS-2 satellite (Reinisch *et al.*, 2007).

Parameterised empirical model (Intercosmos-19)

Pulinets *et al.* (2002) from the Institute of Terrestrial Magnetism, Ionosphere and Radio wave Propagation of the Russian Academy of Sciences (IZMIRAN) in Russia used topside sounder data from the Intercosmos-19 satellite to develop a global empirical representation of the topside ionosphere electron density distribution, during disturbed conditions. They approximated the electron density profile by an Epstein function with an altitude dependent F2-layer thickness parameter. The scale height is described by the following equation:

$$H(h) = H_0 + k(h - h_m) \tag{2.9}$$

Depuev and Pulinets (2004) applied the same technique based on a larger database of topside sounder profiles from the Intercosmos-19 satellite to develop a global empirical model of the topside electron density. They subdivided the model coefficients by the longitudinal sectors and showed the importance of the longitudinal variation of the scale height for the global modelling.

Topside model based on GPS-TEC measurements

The group at the Royal Meteorological Institute of Belgium (Stankov and Muhtarov, 2001; Stankov *et al.*, 2003, 2002) developed a method to reconstruct the vertical distribution of the electron density in the topside ionosphere and plasmasphere using a combination of GPS-TEC measurements, ionosonde measurements and empirically obtained O^+-H^+ ion transition height values. This technique involves the construction of a system of equations based on the different topside profiler functions: Chapman, Epstein and exponential. The system of equations includes a representation of the principle of plasma quasi-neutrality at the F2-peak height, a representation of the fact that the densities of the O^+ and H^+ ions are equal at the UTH (Stankov and Muhtarov, 2001), and the fact that the topside TEC is given by the difference between the GPS-TEC and the bottomside ionosonde TEC. The solution of such a system of equations provides the unknown topside ion scale heights which are then used to reconstruct an unique electron density profile for a given location and time of measurement (Stankov *et al.*, 2003). This technique is discussed in detail in chapter five where the technique is used to create a representation of the topside over South Africa using the co-located GPS and ionosonde measurements from the Grahamstown station.

Topside Sounder Model (TSM)

Kutiev *et al.* (2006) developed a Topside Sounder Model (TSM) that calculates the topside vertical scale height as a function of month of the year, local time, geomagnetic latitude, solar flux $F_{10.7}$ and K_p index. The model is based on topside sounder electron density profiles from the Intercosmos-19 and ISIS satellites. To develop the model, they used a multi-variable polynomial constructed from Chebyshev and trigonometric base-functions, which are fitted to the data in a 5-dimensional space to describe the vertical plasma scale height. From each individual electron density profile, they extract the vertical O^+ scale height (VSH), defined as the lowest gradient of the measured electron density profiles, and the O^+-H^+ upper transition height (UTH), defined as the height at which the upward extrapolated O^+ density becomes half the measured electron density. They proposed a Topside Sounder Model Profiler (TSMP) which provides electron density profiles for three different analytical shapes: Sec-squared, α-Chapman and exponential functions. The TSMP also incorporates the TSM and uses the model quantities as anchor points to reconstruct the topside electron density profile.

2.2.3 Modelling the topside: Theoretical approach

Theoretical or first principles models can be used to represent the behaviour of the ionosphere. This approach involves modelling the physical and chemical processes that are believed to be the major drivers of the distribution of the electron density in the given ionospheric region. For example, the continuity, momentum and energy equations for the electrons and ions are solved as a function of altitude and/or along magnetic field lines to determine the electron density (Schunk and Sojka, 1992). The assumption made is that the calculated electron densities actually describe the variation in the real ionosphere.

Theoretical models have been developed based on the principles described above. For example, Titheridge (1972) developed a procedure for calculating the electron density profile in the topside ionosphere through the "diffusive equilibrium" approach. In this approach, the topside ionosphere is assumed to be composed of the major ions, O^+ and H^+. He simplified the diffusive equilibrium equations to have the scale height given by

$$H_j = -\frac{k_b T_i}{m_j g}\left(1 - \frac{\bar{m} T_e}{m_j T_t}\right)^{-1}$$

where k_b is the Boltzmann constant, T_i is the ion temperature, m_j is the jth ion mass, g is the gravitational acceleration, $\bar{m}$ is the mean ion mass, T_e is the electron temperature and the total temperature is $T_t = T_i + T_e$. Given the number density n_{jo} at the height

h_o, the density n_j at some greater height $h_o + \triangle h$ is calculated from

$$n_j T_t = n_{jo} T_{to} exp(- \triangle h / H_j)$$

This form is suitable for numerical integration, such that, in a series of small altitude "steps", the density of a given ion species at greater height can be calculated if its density at a certain base height is known (Webb and Essex, 2000).

Webb and Essex (2000) used this approach to develop an ionosphere/plasmasphere model modified in such a way that:

(i) Chemical equilibrium is used to model the H^+ ions at lower altitudes and then diffusive equilibrium for the upper ionosphere and plasmasphere.

(ii) The diffusive profiles are calculated along the field lines rather than simply in the vertical direction because the plasma is constrained to move along magnetic field lines.

None of the models that apply the theoretical approach were available to the author at the time of preparation of this thesis, and therefore no analysis was done to assess the applicability of this approach. Efforts to apply this approach to create a model of the topside ionosphere suitable for the South African region were made, and more details are presented in chapter four.

2.2.4 Evaluating the available topside ionosphere models

Efforts were made to access the existing topside ionospheric models and compare their results with measured data in order to determine how well they represent the topside ionosphere over the South African region. Only a few of the models described above were available in the form of a computer program during the preparation of this thesis. These models are presented in this section for comparison with measured electron density profiles for the South African region, and they include the IRI-2007 and NeQuick models.

ISIS-2 topside sounder data processed using the TOPside Ionogram Scaler with True height Algorithm (TOPIST) program, and available through the online archive of the NSSDC at http://nssdcftp.gsfc.nasa.gov/spacecraft_data/isis/topside_sounder/, was used for comparison with the model predictions. The data are described in more detail in section 3.2. Analysis of the model performance involved comparing the electron density profiles predicted by the model with the measured electron density profiles for different

geophysical conditions. However, the measured data is fragmented within the space of the parameters to be studied, such as time of day, season and location. Table 2.2 shows all the available electron density profiles for the South African region within the ISIS-2 (TOPIST) data set (described in section 3.2), and how they are distributed over the 4 years for which data is available. Only a few selected electron density profiles scattered between 1972 and 1975 shown in the table, are available for the South African region. Most of the profiles were for 1973, and coverage is highly non-uniform.

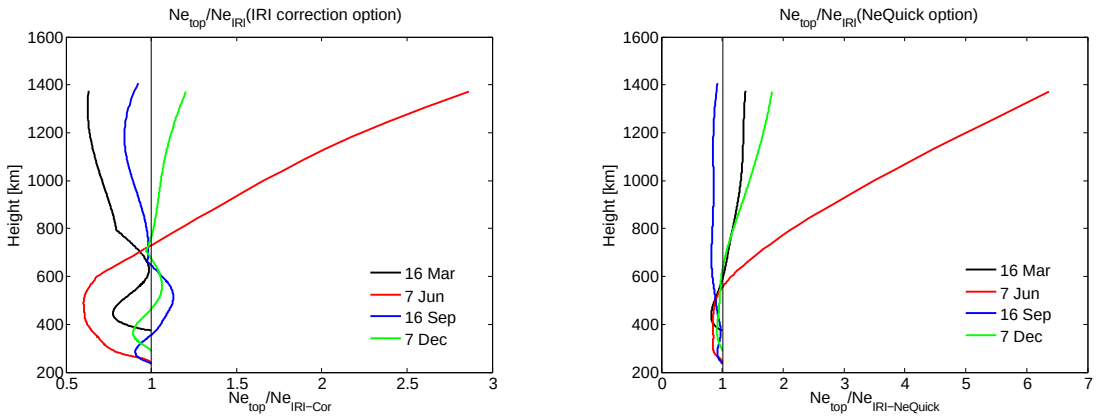
Table 2.2: Available topside sounder electron density profiles from the ISIS 2 satellite. The table indicates the number of electron density profiles available for each month (Sibanda and McKinnell, 2009b).

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1972	-	-	-	-	-	-	-	19	-	-	-	-
1973	-	2	184	54	-	121	1	1	312	-	-	148
1974	17	-	-	-	-	-	-	-	-	-	-	-
1975	-	-	-	-	-	85	-	1	52	-	-	1

Data for 1973 was used for this analysis since it provided fairer coverage of the year with at least each season represented. Sample profiles representing the different seasons were chosen from the database, and the corresponding IRI-2007 and NeQuick model profiles were generated. Both the IRI-2007 and NeQuick models were updated with the measured peak parameters, the critical frequency of the F2-layer ($foF2$) and the maximum height of the F2-layer (h_mF2) from the measured topside profiles, in order to adapt the predicted profiles to the ionospheric conditions of the measured profiles, and concentrate the analysis on how the model reproduces the shape of the profile.

Sample topside sounding data from satellite orbits over the South African region during the daytime were used for comparison with the model results. This sample data occurred on 16 March 1973, 07 June 1973, 16 September 1973 and 07 December 1973, providing profiles which all fell at about the same latitude of around 25.0° - 26.9°S. The data sample was chosen such that the four different seasons were each represented by a day, and individual profiles were selected from those days corresponding to the daytime sector (08:00 to 14:00 LT). Each of the profiles used occurred at a different hour due to the difficulty in obtaining satellite data from the same hour covering all seasons. The ratio between the topside sounder electron density and IRI-2007 electron density was computed for each profile pair, and the results are shown in Figure 2.3.

On the basis of the results shown in Figure 2.3, the IRI-2007 correction factor option



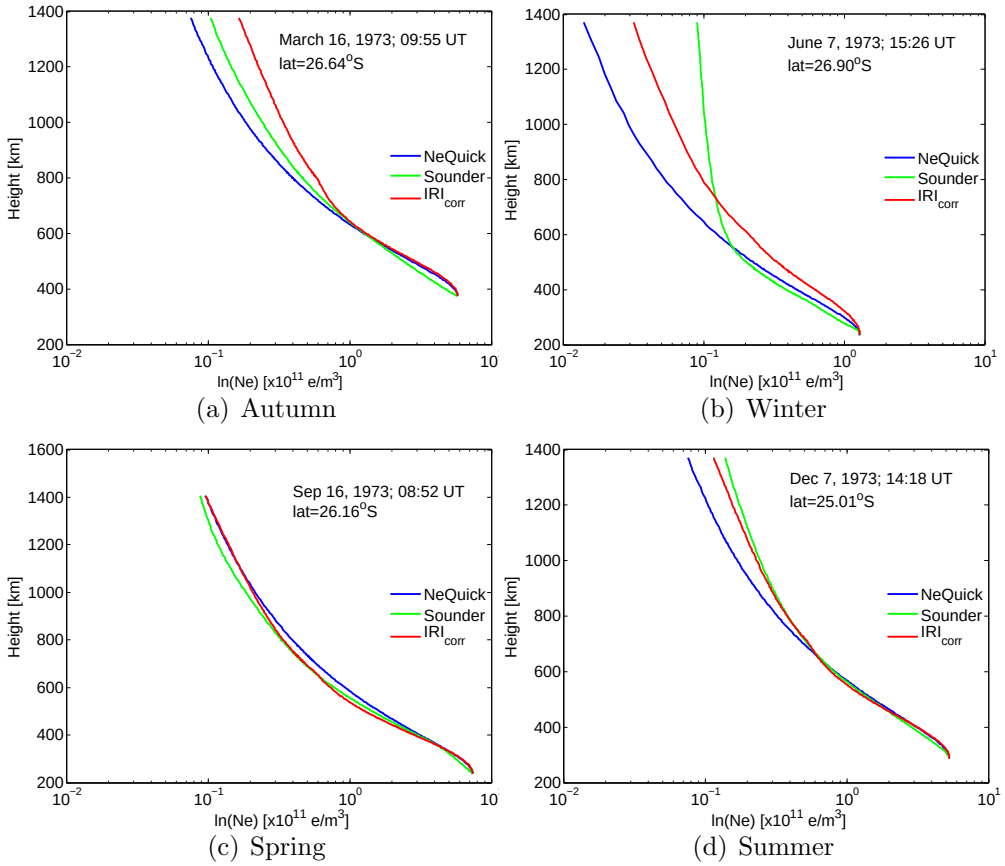
(a) Correction factor option (b) NeQuick option

Fig. 2.3: The ratio between the ISIS-2 topside sounder electron density profiles and model results for the sample moments on 16 March, 07 June, 16 September and 07 December 1973. (a) is the IRI-2007 correction factor option and (b) is the NeQuick option within the IRI model. The local times for the profiles range from 08:00 to 14:00, with each profile at its own hour within that local time range.

model shows greater deviations from the measured data. Although a considerable improvement can be seen when compared with the IRI-2001 version shown by Sibanda and McKinnell (2009a), the IRI-2007 model still slightly overestimates the measured profiles at higher altitudes. The NeQuick model, on the other hand, shows good agreement with the measured data especially at lower altitudes. The ratios of the model and the measurements fall very close to 1 for three out of the four sample profiles used.

Figure 2.4 shows the comparison between predicted electron density profiles and the measured electron density profiles for each of the four sample topside sounder profiles considered, and shows how each individual profile pair compares. The electron density in the profiles shown in Figure 2.4 is plotted on a log scale in order to show a clearer comparison.

These results show, in general, how both the NeQuick and the IRI-2007 models represent the variation of the electron density as a function of altitude. The sample profiles used were chosen to illustrate the ability to predict during different seasons. Both models follow the shape of the measured profile but show considerable deviations in the upper part of the profiles in three out of the four cases shown. The NeQuick model fits well to the measured data at higher altitudes but overestimates the measurements at lower altitudes. The results, to a good extent, are also reflective of the local time variation.



(a) Autumn (b) Winter

(c) Spring (d) Summer

Fig. 2.4: Sample ISIS-2 electron density profiles compared with the corresponding profiles computed with the IRI-2007 and NeQuick models for the dates sampled. These results were published in Sibanda and McKinnell (2009b)

2.3 Summary

The chapter highlighted various aspects of the ionosphere in general and the topside ionosphere in particular, and also highlighted the many efforts made to model this part of the ionosphere, especially the altitude variation of the electron density. Various approaches to topside modelling have been identified and described and a few existing models have been tested to see how they perform over the South African region in order to identify the approach to follow for the South African topside ionospheric modelling efforts. Each of the methods discussed has advantages as well as limitations. Chapters three, four and five present the attempts made to create a representation of the topside electron density distribution over the South African region using some of the techniques described above.

Chapter 3

Empirical modelling based on topside sounder data

3.1 Introduction

Section 2.2.2 provided an overview of the progress in topside ionospheric Ne height profile modelling during the last few years. It also highlighted some major aspects associated with various topside modelling efforts around the world. Emphasis was placed on pointing out the techniques applied in the models that have been established. Those models that were available in computer code or as program software were compared with measured data for the South African region to determine how accurately they represent this region.

Many of the past topside ionospheric studies have been undertaken on a global scale. In contrast, this study is focused in particular on the South African region. Currently the most widely used ionospheric model to predict the topside Ne profile in the South African region is the IRI model. Ground-based ionosondes also provide an estimation of the topside Ne profile based on bottomside ionosphere measurements (Huang and Reinisch, 2001; Reinisch and Huang, 2001). Comparison between the IRI model results and the measured data (section 2.2.4) revealed that, while significant improvements have been made in the latest version of the IRI (IRI-2007), there are still some discrepancies at higher altitudes. A significant number of the studies used available topside sounder measurements to develop models of the topside ionosphere, for example (Bent *et al.*, 1972; Depuev and Pulinets, 2004; Kutiev *et al.*, 2006; Radicella and Zhang, 1995; Reinisch *et al.*, 2007). This chapter presents:

- a comprehensive study of the Alouette and ISIS topside sounder data sets, with a specific emphasis on how the South African region is covered within the databases.
- the attempts made to use the Artificial Neural Network approach to establish an empirical model of the topside ionosphere based on the available topside sounder data. The initial intent was to use the data for the South African region and

develop a regional model. However, the data is unevenly sampled with respect to time and geographical location, and coverage of the South African region is sparse. Limiting the analysis to this region imposed further time coverage limits on the data. This study aims to take advantage of the ability of neural networks to solve complex problems such as the interpolation of unbinned poorly sampled and/or sparse data with the hope that this approach would improve the representation of the global topside, and in turn that of the South African region.

3.2 Database and modelling limitations

This study used the Alouette and ISIS topside Ne profiles archived at NASA's National Space Science Data Centre (NSSDC) in Greenbelt, Maryland. This section provides a detailed discussion of this database. Of the several millions of ionograms recorded by the Alouette and ISIS topside sounder missions, only a few percent were digitised and processed into Ne profiles (Bilitza *et al.*, 2004). This is because (i) the data was recorded in analog form and had to be digitised, which was not an easy process and (ii) the early efforts of processing the digitised ionograms to obtain Ne profiles involved a long and slow process of manually scaling the ionograms. The processed data sets extend from September 1962 (the launch date of Alouette-1 satellite) to 1984. Table 3.1 provides some of the specifics of the data sets indicating the number of Ne profiles available from each of the data sets and the years covered in the processed data.

Table 3.1: Characteristics of the processed Alouette and ISIS data sets

Satellite	**Launch Date**	**Orbit Altitude [km]**	**Available dates for the Processed Data**	**No. of Profiles**
Alouette-1a	1962-09-29	1000	1962-Sep-30 - 1963-Mar-23	15,706
Alouette-1b	1962-09-29	1000	1962-Sep-29 - 1966-Mar-30	43,614
Alouette-1c	1962-09-29	1000	1962-Nov-19 - 1971-Dec-16	26,452
Alouette-2	1965-11-29	500-3000	1965-Dec-15 - 1972-Jun-10	9,301
ISIS-1	1969-01-30	500-3500	1969-Feb-02 - 1971-Mar-28	38,953
ISIS-2	1971-04-01	1400	1971-Apr-08 - 1979-Aug-27	42,596
ISIS-2 (TOPIST)	1971-04-01	1400	1971 - 1984	43,759

The Alouette and ISIS topside sounders provided a fair global mapping of the topside ionosphere covering more than a complete solar cycle. At the NSSDC the analog topside sounder recordings were digitised into ionograms and further scaled and inverted into Ne profiles based on the inversion program of Jackson (1969). They were then converted into a common ASCII data format and made available online from NSSDC's anonymous ftp

site at ftp://nssdcftp.gsfc.nasa.gov/spacecraft_data/. Currently this database consists of the first six data sets listed in table 3.1 and contains a total of about 176,622 profiles. The availability of these data sets stimulated many new topside ionospheric modelling efforts, (Bent *et al.*, 1972; Kutiev *et al.*, 2006; Radicella and Zhang, 1995; Reinisch *et al.*, 2007), and improved the understanding of this critical region of the Earth's upper atmosphere.

Additional and more recent data restoration efforts to process more digitised ionograms into Ne profiles resulted in another data set of ISIS-2 Ne profiles, processed using the TOPside Ionogram Scaler with True height Algorithm (TOPIST) program. This data set extends from 1971 to 1984 and consists of about 43,759 Ne profiles (Bilitza *et al.*, 2004; Huang *et al.*, 2001). TOPIST is based on the techniques of automated scaling procedures developed in the 1980s (Huang and Reinisch, 1982) and have been used successfully with ground-based ionosonde measurements (Reinisch *et al.*, 2001).

The ISIS-2 (TOPIST) data set is available from the online archive of the NSSDC at http://nssdcftp.gsfc.nasa.gov/spacecraft_data/isis/topside_sounder/. This database complements the manually scaled database described above to improve coverage of the relevant geophysical conditions (time, season, geographical location and solar activity).

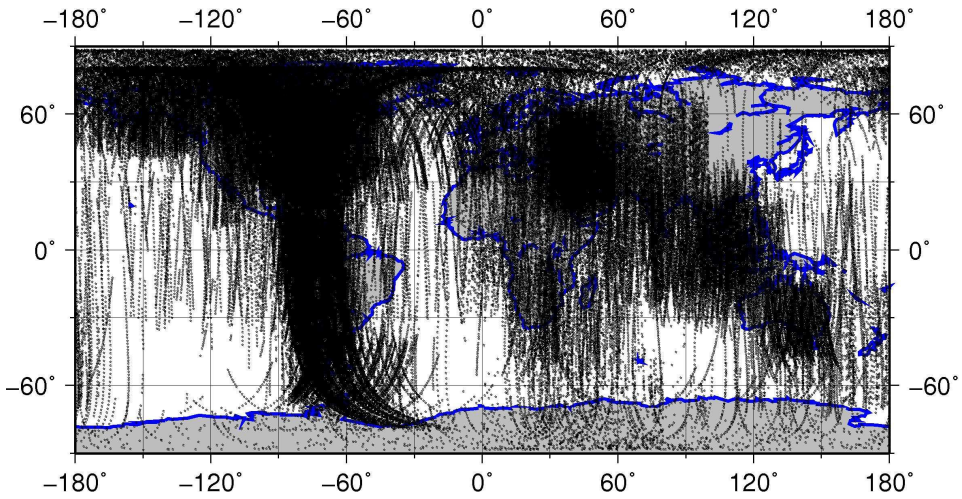
3.2.1 Coverage of the South African region

The main problems encountered with the Alouette and ISIS topside sounding missions has been the quantity of the processed data. Coverage depended on the distribution of the telemetry stations because the satellites either had limited on-board recording capabilities (such as ISIS) or had none at all (Alouette). Data were primarily recorded within the viewing area of the telemetry stations.

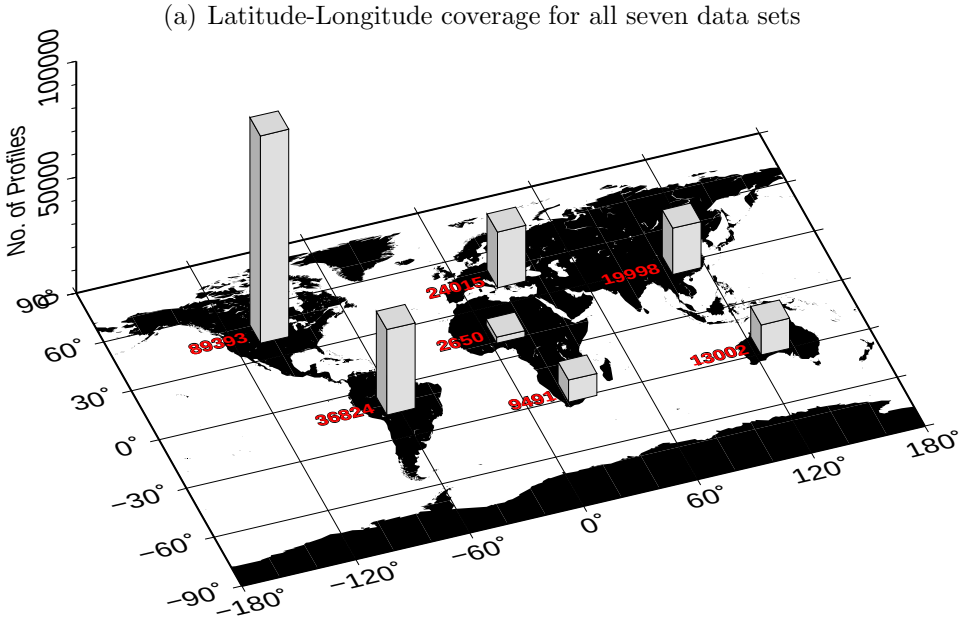
In addition, due to the cumbersome manual scaling process involved (Bilitza *et al.*, 2004), only a small percentage of the data was processed from the original ionograms on microfilm to Ne profiles, which is the parameter of greatest interest for topside ionospheric studies. Most of the data extending up to the time the ISIS-1 & -2 satellites shut down in January 1990 were never processed into Ne profiles. Due to these factors and the fact that, satellites don't make systematic and repeatable temporal and/or spatial observations since the satellite is constantly in motion, the measurements are unevenly sampled and relatively sparse (Huang *et al.*, 2002). This has posed the greatest challenge to the studies of the topside ionosphere.

Figure 3.1 shows the global coverage provided by the combined Alouette/ISIS data sets

listed in table 3.1. The first plot (Figure 3.1(a)) shows the longitude zones covered in the data, which illustrates a strong concentration on the American sector. The distribution of the Ne profiles over the different regions is indicated in the second plot (Figure 3.1(b)). Coverage of the South African region is minimal within this database with only 9491 profiles for the South African region from a total of 211,139 profiles, representing about 4.5% of the whole data set.



(a) Latitude-Longitude coverage for all seven data sets

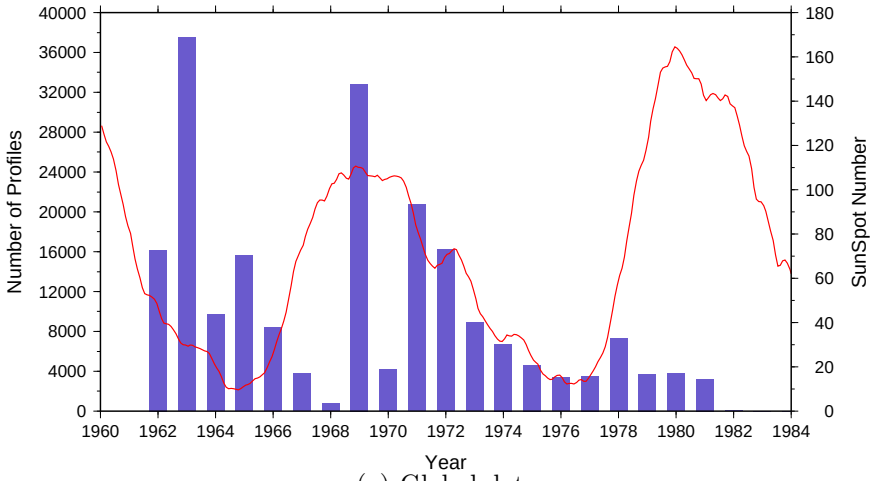


(b) Distribution of the Ne profiles into the different regions

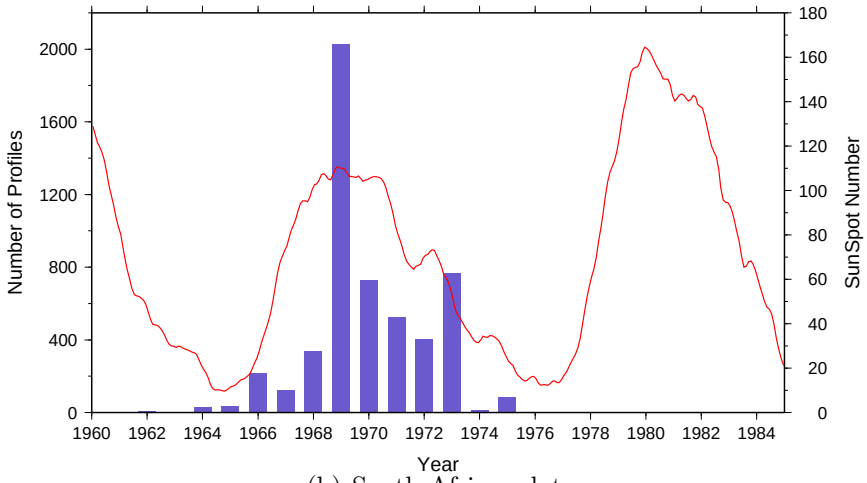
Fig. 3.1: Geographical coverage of the processed Ne profiles for the seven data sets

Figure 3.2 shows the distribution of the data over the years extending from 1962 to 1984 in relation to the solar cycle variation. The data collection period covers more than one entire solar cycle. While on the global scale the data is distributed fairly evenly over all

the years, the distribution for the South African region (Figure 3.2(b)) is limited to the period between 1966 and 1972, a period of mostly high solar activity.



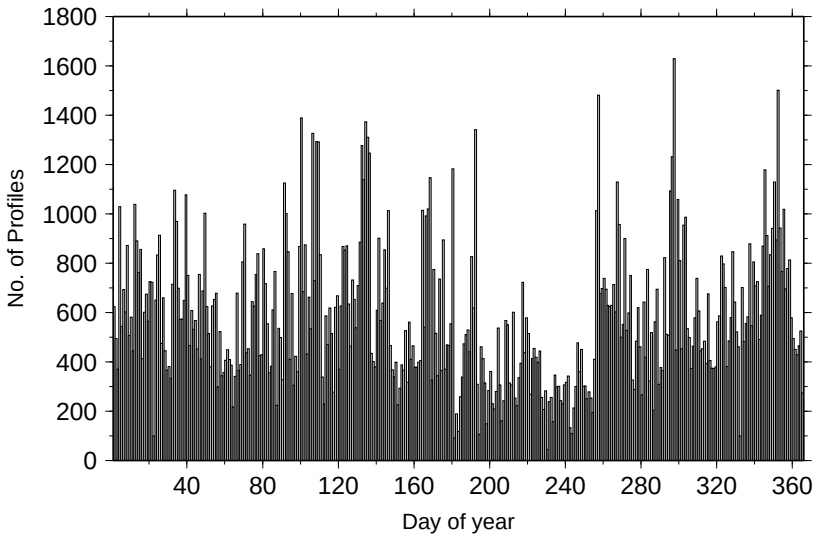
(a) Global data



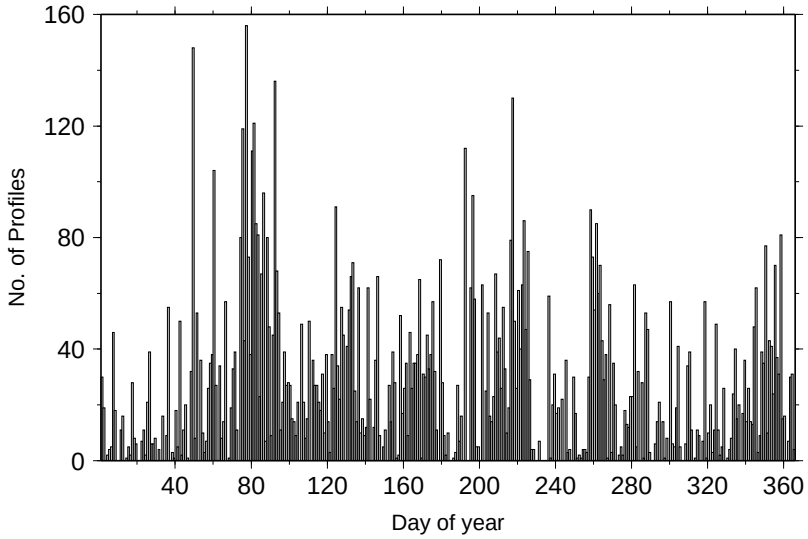
(b) South African data

Fig. 3.2: Distribution of the processed Ne profiles over the years from 1962 to 1979 in relation to the sunspot cycle

The day of year coverage plot for the global and South African data in Figure 3.3 shows fairly uniform coverage. This assures proper statistical sufficiency of the data set for the studies of seasonal characteristics of the topside Ne variations. The local time coverage is quite uniform. The spatial and local time coverage can be considered to be statistically reliable.



(a) Day of year coverage for the global data



(b) Day of year coverage for the South African region

Fig. 3.3: Day-of-year distribution for a) all available data and b) the South African region.

3.3 Modelling based on the topside sounder data

Proper understanding of the dynamic structure of the Earth's ionosphere is of fundamental practical importance because of the essential part it plays in telecommunication and navigation systems where radio waves are required to pass through the ionosphere. In particular, characterisation of the vertical distribution of the Ne in the ionosphere is of ever increasing importance because most of the ionospheric effects on radio signals are directly proportional to the Ne. Modelling the Ne height profile over the whole altitude range for all geographic positions and time sectors is critical in the development

of practical schemes for providing reliable prediction of HF propagation. The ionosphere is highly non-linear and its structure varies significantly under various geophysical conditions such as local time, geographical location and solar activity level (see section 2.1.1), and thus, modelling it is not always easy. In the following section, the attempts made to create a model of the topside ionosphere based on topside sounder data using the technique of Artificial Neural Networks (ANNs) is outlined.

3.3.1 Using Artificial Neural Networks

Due to the complexity of its structure, successful modelling of the ionosphere involves complex procedures which require proper representation of all the factors that influence its behaviour. Studies of the topside ionosphere under natural conditions are associated with many disadvantages and limitations owing to the difficulty of measuring that region. Due to the lack of measurements, many external and irregular factors that exert an influence on the behaviour of the topside ionosphere are often not fully known. The vertical structure of the topside ionospheric Ne in particular depends on many variables and the relationships between these multiple variables and the Ne distribution are highly non-linear or chaotic in nature, making it difficult to capture them mathematically. Determining these relationships is a type of function approximation problem that Artificial Neural Networks (ANNs) can be used to solve, typically using the method of multi-layer perception.

In chapter two (section 2.2.2.2), an overview is given of the existing empirical models of the topside ionosphere emphasising the techniques applied in developing these models. These approaches mainly involved binning the data and assuming physics-based analytical function whose coefficients are obtained from the data. In this chapter, an attempt is made to use an approach involving the ANN technique to establish an empirical model of the topside ionosphere. ANNs are currently actively used in various fields including business, medicine, satellite imagery, meteorology and many more for such applications as time series prediction (Fausett, 1994), pattern classification and non-linear function approximation. They have also been used in ionospheric research and proved to be a powerful tool for forecasting ionospheric behaviour (McKinnell and Poole, 2004). The next section presents the basics of ANNs and how they have been used in ionospheric characterisation. This chapter is focused on using the ANNs rather than studying them, and therefore, readers are referred to Fausett (1994) and Mehrotra *et al.* (1996) for more details on ANNs.

3.3.1.1 Artificial Neural Networks - description

ANNs are computer programs which, when presented with a set of multi-dimensional input data, the program learns to map the inputs onto a set of appropriate measured output parameters. To use the ANN programs, one needs to build an architecture (selecting the appropriate learning function and choosing the right number of input and hidden nodes), which best suits one's requirements. For example, Williscroft and Poole (1996) tried various architectures and found that a feedforward backpropagation learning algorithm was optimum for ionospheric applications. Generally, in ionospheric research, the appropriate ANN architecture should be able to capture the relationship between the set of input observations and the desired set of outputs. Furthermore, since the behaviour of the ionosphere is extremely complex, the appropriate ANN architecture is one that has the ability to find approximate solutions to problems and processes with non-deterministic behaviour. The backpropagation algorithm has this ability, and is therefore the standard algorithm used in this study.

Generally, a multi-layered ANN consists of an input layer, one or more hidden layers of non-linearly activating nodes and an output layer. A typical artificial neuron and the modelling of a multi-layered ANN are illustrated in Figure 3.4.

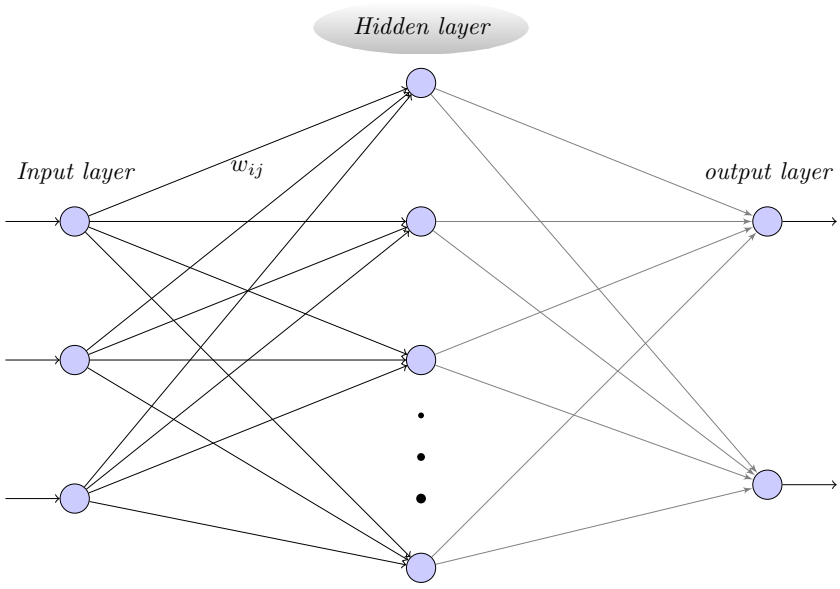


Fig. 3.4: An illustration of an artificial neuron and the modelling of a multi-layered neural network

Each node in one layer connects to every other node in the following layer with a randomly chosen weight w_{ij}. The data are randomly ordered and presented to the network iteratively for processing. The ANN learns to identify the relationship between the input parameters and the known output parameters by changing connection weights,

based on the amount of error in the output compared to the expected measured result after each piece of data is processed (Fausett, 1994). The weights are changed in such a way as to minimise the difference between the output produced and the measured output, and this is repeated until the root mean square error (MSE) between the measured and predicted outputs on the testing set is stabilised (Mehrotra *et al.*, 1996). Training the ANN requires a large database of measurements for several parameters that are used as input and output vectors, collected over a period of time, from which the ANN can learn to capture the underlying functional relationships between the input and output variables.

ANNs have been used widely in ionospheric research in South Africa and have proved to be very reliable in modelling the complex dynamics of various ionospheric parameters (e.g characterising bottomside Ne profile (McKinnell, 2002; McKinnell and Poole, 2004), predicting the occurrence of spread F (Paradza, 2008) and Total Electron Content (TEC) prediction (Habarulema *et al.*, 2007)).

3.4 Preparing the data for training

The goal as outlined in the introduction to this chapter is to create a model of the topside ionosphere Ne as a function of height, based on the information obtained from the measured topside sounder data. Therefore, the seven topside sounder data sets archived at the National Space Science Data Center (NSSDC), Greenbelt, MA, and listed in table 3.1, were processed and prepared for feeding into an ANN to create the model. This involved:

- Checking the integrity of the data in order to exclude the erroneous data points.
- Defining the model variables, that is, the desired outputs and determining the relevant input parameters.
- Putting the data into the form accepted by the neural network. This involves categorising the data into the training, validating and testing sets.

3.4.1 Data integrity

Many of the topside sounder Ne profiles are terminated at the last observed frequency on the topside ionogram, which is less than the F2-layer critical frequency due to the limitations in the inversion procedure used to convert the ionogram traces to Ne profiles (Bilitza and Williamson, 2000). However, the use of the data in this study required an estimation of the F2-peak parameters ($foF2$ and h_mF2) from the profiles. To solve this

problem, the profiles were extrapolated down to the F2-peak by fitting an α-Chapman function to the lower end of the profile, using a non-linear least squares approach to find the best F2-peak parameters ($foF2$ and h_mF2) and the scale height H_m. Figure 3.5 shows an example of a profile with an α-Chapman layer fitted to the bottom end.

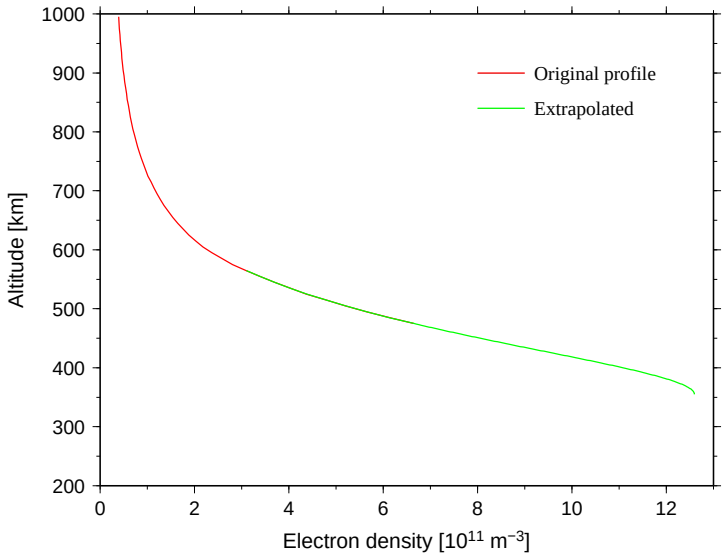


Fig. 3.5: ISIS 2 topside Ne profile with an α-Chapman function fitted to the bottom end of the profile.

This procedure was performed on all the Ne profiles used. An α-Chapman function is a very good fit to the measured topside sounder profiles only for the first few hundred kilometers ($\approx$ 200) (Reinisch *et al.*, 2007) above the F2-peak. Thus care was taken in choosing how far up the profile to fit the function. Major differences were noted among the data sets. Some of the profiles, mainly those from the Alouette-1a and ISIS-1 for example, were terminated at about 500 km even in instances where it is expected that the peak should be much lower. These profiles caused difficulties in the processing because the fitting procedure could not converge and the F2-peak parameters could not be determined. Such profiles were not included in the analysis. In the case of the ISIS-2(TOPIST) data set, the calculated F2-peak parameters were close to the values given in the data and the F2-peak values were determined with greater ease.

3.4.2 Identifying the model output variables

The ultimate goal is to create an empirical model that will output the Ne as a function of height. This is a complicated task that requires taking into account all the different factors that play a role in its variation. Therefore, direct modelling of the Ne distribution is not easy. In this study, the densities of the individual ions that have a significant influence on the distribution of the Ne in the topside ionosphere are modeled, from which the Ne is derived. This allows the inclusion of other important profile shape

factors which simplifies the problem. For example, if the UTH is known, it can provide information about the relative composition of the ions, which can be used with other parameters to reconstruct the Ne profile. In this regard, three parameters (the O^+ and H^+ vertical scale heights and the UTH), are derived from the measured Ne profiles. The derivation of these parameters is based on the key assumptions that the height profile of the topside individual ions exponentially decrease from the F2-peak height, and that O^+ and H^+ are the major ions in the topside ionosphere. These parameters can then be used to reconstruct the vertical distribution of the Ne in the topside ionosphere following the procedure described in section 3.7.

The upper transition height (UTH) is an important characteristic of the ionosphere. It serves as a boundary between the O^+ dominated topside ionosphere and the H^+ dominated plasmasphere. The physical significance of this height is that the densities of the major ion species O^+ and H^+ are equal, and the slope of the Ne profile changes rapidly within the vicinity of this transition region due to the different gradients of the H^+ and O^+ ion profiles. If known, it can serve as a base for finding the relative quantity of H^+ and O^+ ions. Several studies have used different approaches to derive the UTH from topside sounder Ne profiles. For example, Titheridge (1976) extracted the UTH from the Alouette-1 topside Ne profiles by fitting the data with a theoretical model that is based on diffusive equilibrium. Using about 60,000 profiles measured between 1962 and 1968, he obtained transition height variations for various seasons and levels of solar activity. Marinov *et al.* (2004) derived the UTH from topside sounder (Alouette-1a, -1b and -1c and ISIS-1 and-2 satellites) Ne profiles by plotting the profiles on a natural logarithmic scale and then fitting a regression line over the points at which the gradients do not exceed the lowest by 30%. They extrapolated the regression line to higher altitudes and defined the transition height as the height at which the Ne gradient yields a density which is one half of the measured Ne.

Marinov *et al.* (2004) used these UTH values to create an empirical representation of the UTH as a function of local time, month of the year, geomagnetic latitude and longitude and solar radio flux (F10.7). Their mathematical formulation was based on the approach that was first developed by Kutiev *et al.* (1984), and further modified (Kutiev *et al.*, 1994) to include more data sets. Other studies have used other satellite data sources to create empirical models of the UTH. Miyazaki (1979) for example, developed a model based on TAIYO satellite data. He obtained a semi-empirical formula for UTH and was able to observe and explain the observed asymmetry of ion transition heights between the Southern and Northern hemispheres. Kutiev *et al.* (1984) used a generalised

multivariable polynomial, containing a system of linearly independent functions in five dimensions, to create a model of the UTH. The input UTH values were derived from the OGO-6 satellite encounters of the UTH. The model provided the UTH as a function of sunspot number, month of the year, local time, dipole latitude and longitude.

3.4.3 Defining the model variables from the Ne profiles

The UTH is defined in terms of the respective vertical scale heights of the oxygen ion (O^+) and the hydrogen ion (H^+) vertical profiles. The Ne profile is plotted on a natural logarithmic scale with the Ne on the x-axis. The vertical scale heights of the O^+ and H^+ are represented by robust regression lines fitted respectively to the bottom end and the top end of the Ne profile.

3.4.3.1 Ion vertical scale heights

O^+ vertical scale height (H_{O^+})

First the height at which the lowest gradient of the measured Ne profile occurs is determined. Below this height, the effects of the F-region recombination processes play a significant role and therefore the gradient increases sharply such that the scale height at the peak itself is infinite. Above this height, on the other hand, the gradient increases with height due to the increase of plasma temperature (Titheridge, 1976). To account for this temperature increase the H_{O^+} is defined from the O^+ gradient in this region (Kutiev *et al.*, 2006). The O^+ gradient is obtained from a robust regression line fitted over the points above the height of lowest gradient such that the gradients do not exceed the lowest by 20%. The values of the H_{O^+} derived ranged between 100 and 200 km.

H^+ vertical scale height (H_{H^+})

As the upper transition region is approached, it is not only the plasma temperature that plays a role in the increase of the Ne gradient, the H^+ density also starts contributing significantly. Although it is difficult to distinguish between H^+ and temperature contributions, it was assumed in this study that H^+ plays a major role higher up above the upper transition height. Therefore, the H_{H^+} is approximated by fitting a robust regression line at the top end of the profile to the points for which the gradients are not more than 20% less than the highest. The H_{H^+} values ranged between 200 and 2000 km.

Robust regression was chosen over the ordinary least squares regression because it is relatively less vulnerable to unusual data, especially around the regions along the profile where the gradient of the curve changes rapidly such as around the F2-peak and the upper transition height. Robust regression is a compromise between deleting these

points, and allowing them to violate the assumptions of ordinary least squares regression. The robust regression technique assures a more stable result and avoids the uncertainties invoked by the data scatter around the curving areas.

Table 3.2: Examples of the ratios H_{H^+}/H_{O^+} on 20 March 2000 at different times of the day

Hour	$\mathbf{H}_{O^+}$	$\mathbf{H}_{H^+}$	$\mathbf{H}_{H^+}/\mathbf{H}_{O^+}$
00h00	181.37	929.38	5.12
12h00	157.18	583.84	3.71
21h00	149.75	891.29	5.95

Theoretically, in diffusive equilibrium along the magnetic field lines assuming an isothermal ionosphere, the ratio of the oxygen and the hydrogen scale heights H_{O^+}/H_{H^+} is equal to their mass ratio ≈16. In this study, e.g table 3.2, the scale height ratios are lower than the theoretically expected ratios. It is worth noting that the scale heights are derived in the vertical direction. Other factors that can affect the theoretical ratio include the fact that the horizontal structure is non uniform, which may disturb the vertical ion distribution.

3.4.3.2 Upper transition height (UTH)

The UTH is then defined as the point at which the O^+ robust regression line intersects with the H^+ robust regression line. Figure 3.6 illustrates the above definition.

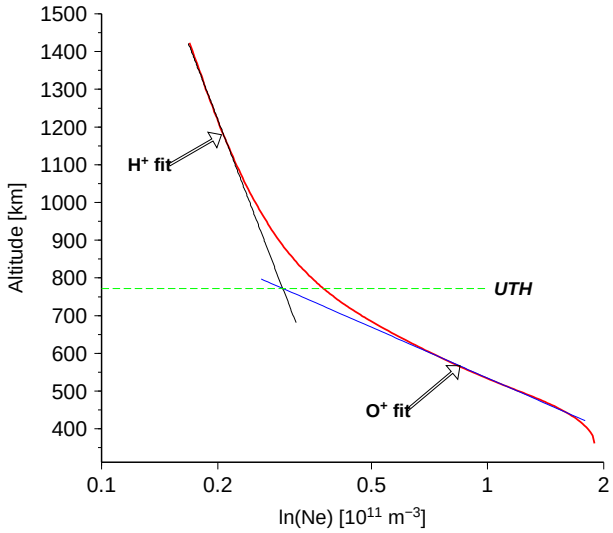


Fig. 3.6: The upper transition height definition; determined by fitting the O^+ regression line at the lower end and the H^+ regression line at the top end of the profile on a logarithmic (ln(Ne)) scale

The measured Ne (red line) plotted on a natural logarithmic scale and the O^+ gradient (blue line) and the H^+ gradient (black line) regression lines are fitted at the bottom

and top ends of the Ne profile respectively. The transition height (dashed green line) is the altitude at which the regression lines intersect. This is an approximation since the transition is gradual and it is impossible to determine the exact position where the profile changes shape by inspection of the profile.

This definition assumes that the O^+ and H^+ are the major ions present in the topside ionsphere, the presence of He^+ is therefore neglected. The Ne is equal to the sum of the ion densities present in the ionosphere, in this case the O^+ and H^+ ions. $Ne = N_{O^+} + N_{H^+}$. The values of the gradients and the upper transition height are extracted from each individual profile and prepared for creating the model.

Both the O^+ and H^+ vertical scale heights are considered to be constant throughout the topside ionosphere. The robust regressions are taken over several Ne values which actually represent scale heights with an average plasma temperature in the respective altitude ranges. The jump from the value of the H_{O^+} to the value of H_{H^+} for a given profile highlights the change in plasma scale height between the lower and the upper part of the topside profile. Theoretically, the plasma scale height must increase with altitude due to the increase in the plasma temperature. This procedure therefore takes into account the increase in plasma scale height with altitude.

3.4.4 Determining the model input parameters

The database also provides other ionospheric parameters corresponding to each Ne profile, which include: local time, geographic latitude and longitude, magnetic inclination (dip) and solar zenith angle at 100 km. These parameters were also collected and used to build the input space for the ANN.

The variation of the ionosphere is influenced by several factors caused by the Sun and its activities (section 2.1.1). Special attention was paid in determining the model parameters to ensure these factors were taken into consideration when choosing the input parameters for the ANN. The diurnal variation was represented by the local time (Hr). The day of year (DoY) was used to capture the ionospheric variation due to the seasonal changes. Ideally, the DoY ranges between 1 and 365. However, the data is fragmented with respect to this parameter with many gaps in the data. Solar activity changes also influence the ionospheric behaviour. As a measure of the solar activity, the 12-month running mean of the sunspot number (Rz12) was used. Rz12 was calculated by averaging the monthly values of the sunspot number centred on the current month. During the period covered by the data (more than a complete solar cycle), Rz12 varies in value from 0 to about 120. The geographic latitude (glat) and longitude (glon) were used to capture the variation

with respect to geographic location. The sampling of the data with respect to these parameters is highly non-uniform as shown in Figure 3.1. Tests to exclude either glat or glon as inputs indicated that both these parameters have significant influence on the model output. The other candidates for the model variables included magnetic inclination (dip) and solar zenith angle (Zenith). The magnetic inclination was included in order to account for the effect of the inclination of the magnetic field on the variation of Ne. This is because the plasma is confined to diffuse along magnetic field lines according to the theory of plasma diffusion, which plays a major role in the distribution of the Ne in the topside ionosphere. Finally, the F2-peak parameters, the maximum height of the F2-layer (h_mF2), and the maximum electron density of the F2-layer ($NmF2$) were also included as inputs to the ANN. As pointed out earlier, the distribution of ionisation in the topside ionosphere is closely linked with the production and loss of ionisation in the bottomside ionosphere, especially in the F-region. Therefore, including h_mF2 and $NmF2$ as inputs ensures that the ANN learns to capture the link between the topside ionosphere and the F2-peak dynamics. This also ensures that the model can be smoothly merged with the bottomside profile.

The daily and annual cycles can cause a problem in the way the ANN learns to recognise this cyclic behaviour. For example, the ANN may not be able to recognise the adjacent years as different cycles. To circumvent this, both the day of year (DoY) and hour (Hr) parameters were split into the sine and cosine components for each parameter (Williscroft and Poole, 1996), such that:

$$\left.\begin{aligned} DoY\_S &= sin\left(\tfrac{2\pi\times DoY}{365.25}\right) \\ DoY\_C &= cos\left(\tfrac{2\pi\times DoY}{365.25}\right) \end{aligned}\right\} \text{ sine and cosine components of DoY}$$

$$\left.\begin{aligned} HrS &= sin\left(\tfrac{2\pi\times Hr}{24}\right) \\ HrC &= cos\left(\tfrac{2\pi\times Hr}{24}\right) \end{aligned}\right\} \text{ sine and cosine components of Hr}$$

This also ensures that the ANN learns to avoid a step jump from the end of one year to the beginning of the following year in the case of the annual cycle. The same applies with the diurnal cycle to avoid a jump between adjacent days.

3.4.5 Training, validating and testing data sets

From each individual Ne profile, values of the UTH, H_{O^+} and H_{H^+} were extracted together with the corresponding values of the input variables, and prepared for training

the ANN. A total of 211139 values were available. The data were arranged into three categories: the training and validating sets which are loaded into the network, and a testing set that is not seen by the network. The training and validating sets were converted into the format accepted by the ANN. Evaluating the performance of the ANN involves testing it on a set of new data not belonging to the training set. Thus, a separate set of data must be set aside as a testing set. The testing data set included: (i) part of the measurements taken in 1978 and 1979 which falls outside the range of the training set to represent the low solar activity period and (ii) part of the measurements taken between 1971 and 1975 which falls within the range of the training set to represent the high solar activity period. The choice of the testing data also allows for testing how well the ANN is able to both interpolate and extrapolate the patterns learned.

3.5 Creating the model and the model architecture

Various combinations of the above ionospheric variables were investigated in order to determine the combination that yielded the most accurate results. The set of input parameters yielding the best approximation to the outputs are: the sine and cosine components of hour (HrS, HrC), the sine and cosine components of the day of year (DoY_S, DoY_C), latitude, longitude, dip, Zenith, Rz12, h_mF2 and $NmF2$.

The architecture of the ANN used had one input layer with eleven input nodes for the eleven parameters respectively, one hidden layer and one output layer. In addition to choosing the right number of input nodes, special consideration must also be taken in choosing the appropriate number of hidden nodes and hidden layers. The number of hidden nodes affects how well the network is able to separate the data. A network consisting of a large number of hidden nodes is compromised in its ability to generalise and therefore cannot perform well on data that was not part of the training (Ajith, 2005). On the other hand, with too few hidden nodes, the network may be unable to learn the relationships amongst the data and the error may fail to fall below an acceptable level. Thus, the selection of the number of hidden nodes is an important decision. Simulations revealed that generally one hidden layer consisting of a number of hidden nodes equal to the number of input nodes + 3 is an optimal architecture for the situation described in this study.

The ANN software used is the Stuttgart Neural Network Simulator (SNNS) version 4.2. The SNNS was developed by the Institute for Parallel and Distributed High Performance Systems, University of Tübingen and the Wilhelm Schickard Institute for Computer Science in Germany, and is freely available at http://www.ra.cs.uni-

tuebingen.de/SNNS/.

Several combinations of the numbers of hidden nodes and learning functions, including the standard back-propagation learning algorithm, were tested. The best performing network architecture consisted of the standard back-propagation learning algorithm with one hidden layer of 14 nodes, along with eleven input nodes and three output nodes representing the predicted parameters, as shown in Figure 3.7.

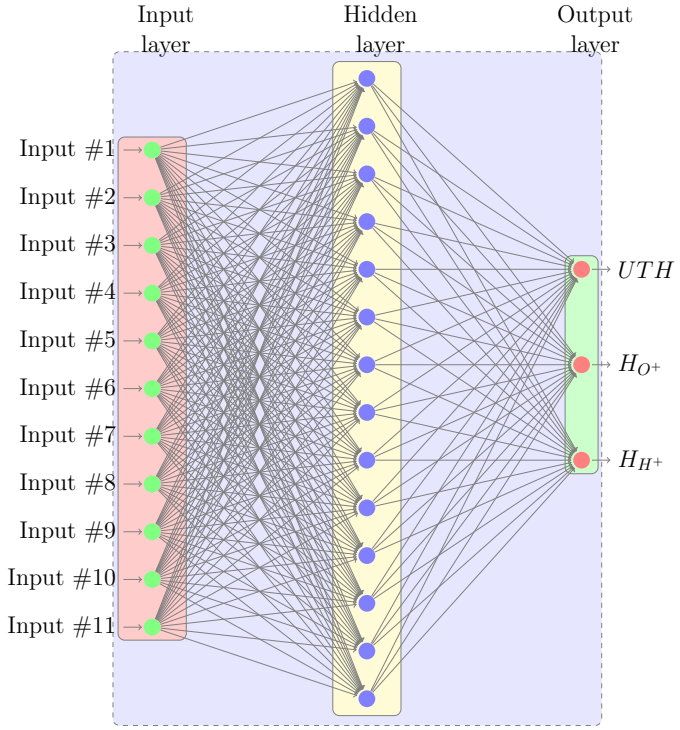


Fig. 3.7: The ANN architecture that gave the best performance. The inputs correspond to DoY_S, DoY_C, HrS, HrC, glon, glat, dip, zenith, Rz12, h_mF2 and $NmF2$. The outputs correspond to H_{O^+}, H_{H^+}, and UTH

3.6 Model results and analysis

This section presents the results and analysis of the ANN model developed. The analysis involves only those results covering the South African region which is the region of interest for this thesis. The model is based on the Alouette and ISIS topside sounder data and predicts the UTH, H_{O^+} and H_{H^+} given the local time, geographic latitude and longitude, magnetic inclination, solar zenith angle, 12-month running mean of the sunspot number (Rz12), h_mF2 and $NmF2$. The model is computed for various ionospheric conditions in order to investigate the dependence of the model variables on various factors including time of day, season and solar activity level. The analysis was carried out for two ionosonde stations in South Africa, Grahamstown (33.3°S, 26.5°E)

and Madimbo (22.4°S, 30.9°E), which are separated in geographic latitude by about 10°.

3.6.1 Data limitations

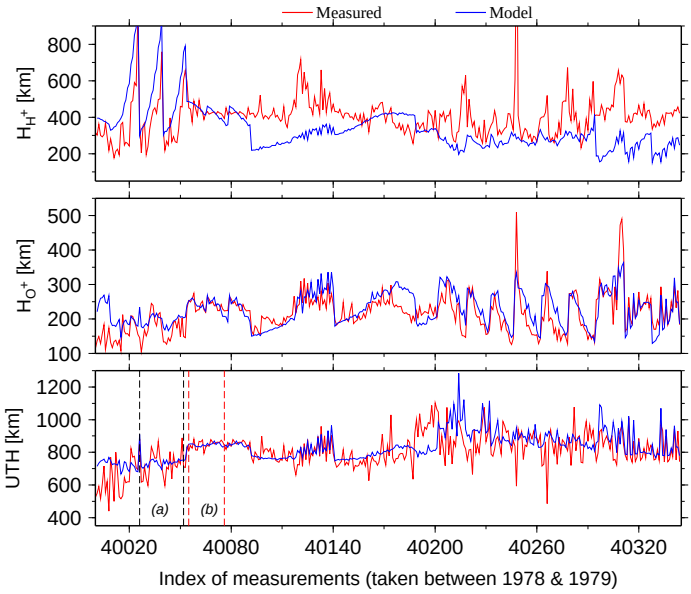
The sampling of the data is highly non-uniform. It is fragmented within the input space particularly in representation of time of day, day of year and geographical location. The satellite data only provides sets of data points along the satellite path, such that for example, only short time measurements of up to three minutes are available for a given location on a particular day. For each day represented, the database does not sample all the hours of the day in a way that will allow for distinguishing the Ne behaviour for different hours of the day. Similarly with the day of year, not all days contain data. It is therefore difficult to make systematic comparisons between the model results and the measured data. The longitude and latitude variations of the model were not investigated in the current study, and thus, only the diurnal, seasonal and solar activity level investigation will be presented.

3.6.2 General trend

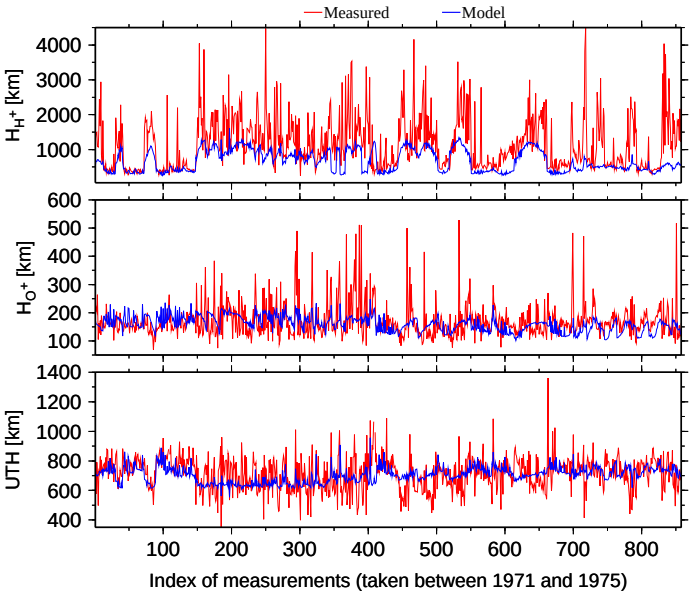
In the first analysis, longitude and latitude considerations were neglected. All the data were combined together according to day of year from the first available day in 1978 to the last available day in 1979, regardless of the geographical location where the measurements were taken. Figure 3.8 is a general trend plot that shows the comparison between the model results and the measured data.

Figure 3.8(a) is a plot of measurements taken between the years 1978 and 1979, which represents a data set that falls outside the range of training data. This illustrates how the ANN was able to extrapolate outside the range of the training set. On the other hand, Figure 3.8(b), is a plot of a selection of data points during the period between 1971 and 1975, which illustrates how well the ANN is able to interpolate. Both data sets were not included in the training set. The x-axis is the index of the data points representing the short time measurements taken at various locations (latitude, longitude) along the satellite path. For each particular day of year represented, there are about six data points taken over a duration of about three minutes, and the hour at which the measurements were taken varies from day to day. The sampling also depended on the profiles that were processed, since only a selection of the ionograms were processed.

It appears from the two plots that the model follows the general trend of the measurements both during high solar activity and low solar activity periods. To expand this analysis and to elaborate the results further, two moments that lay within the South African region labeled (a) and (b) in Figure 3.8(a) are selected and plotted as shown



(a)



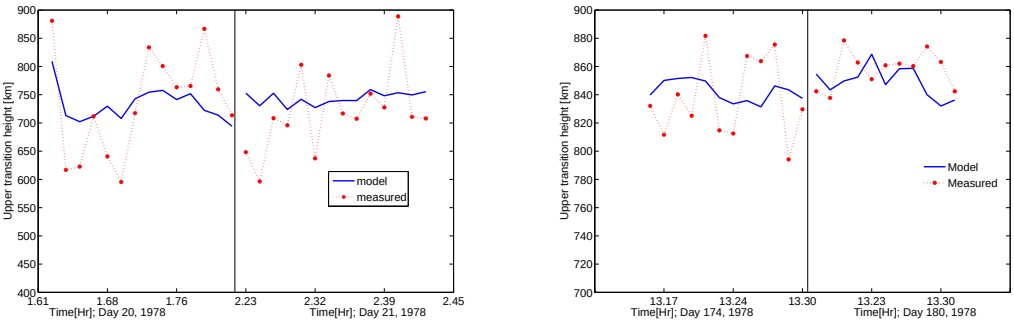
(b)

Fig. 3.8: A general trend plot, comparing the measured and model values of the UTH (bottom panel), H_{O^+} (middle panel) and H_{H^+} (top panel). On the x-axis is the index of: (a) the data points taken during a low solar activity period between 1978 and 1979 and (b) a selection of measurements taken during a high solar activity period between 1971 and 1975.

in Figure 3.9. The moments correspond to days 20-21 (Figure 3.9(a)) for a nighttime example and days 174 and 180 (Figure 3.9(b)) for a daytime example.

The measurements occurred during different seasons. This comparison shows that generally, the model predictions follow a good fit through the measured values which

means that the model neither underestimates nor overestimates the measurements.



(a) Nighttime trend comparison between model & measurements

(b) Daytime model/measurement trend

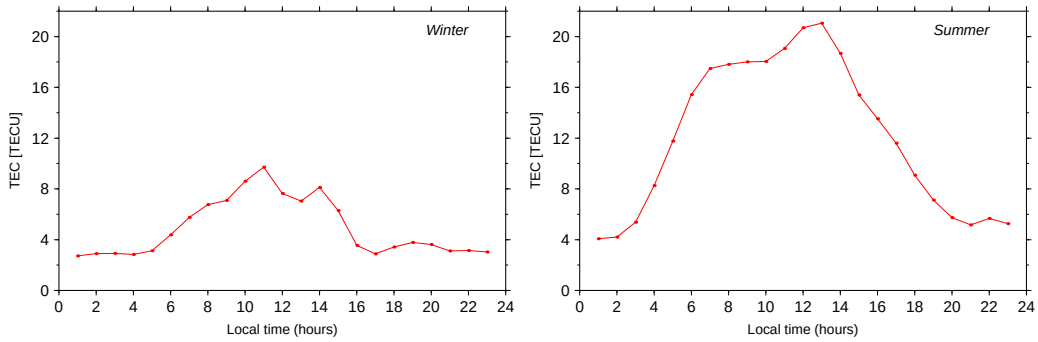
Fig. 3.9: The two snapshot measurements within the results presented in Figure 3.9(a) corresponds to the snapshot measurements taken on days 20 and 21 in 1978 labeled (a) on Figure 3.8 and Figure 3.9(b) corresponds to days 174 and 180 of the same year marked (b) on Figure 3.8. The data does not show a complete day, only sets of data points for the duration of the satellite pass over the region of interest.

It is important to note, however, that the above analysis only provides the general view of how the model is able to follow the trend of the random measurements. Very little information can be drawn from the two plots regarding how well the model follows the expected diurnal and seasonal trends.

3.6.3 Upper Transition Height variation

As noted before, the data is not uniformly distributed with respect to any of the parameters required for topside modelling. For the purpose of analysing the model performance in order to determine the kind of daily, seasonal, and other such cyclic variations that the model follows, it was necessary to consider the daily, seasonal and solar activity behaviour of other ionospheric parameters and analyse the observed variations of the UTH based on this. The dynamics of the topside ionosphere are closely coupled with those of the bottomside ionosphere especially in the F-region, although the behaviour in the topside ionosphere is rather different from that in the bottomside ionosphere. The UTH variations were analysed with respect to the variations of the total electron content (GPS-TEC) obtained over Grahamstown, South Africa without necessarily expecting them to vary in the same way. GPS-TEC varies diurnally with higher values during the daytime and lower values at night.

Figure 3.10 shows examples of TEC diurnal variation for a winter day (Figure 3.10(a)) and a summer day (Figure 3.10(b)). The magnitude of the TEC values over Grahamstown (a mid-latitude station) are higher in summer than in winter.



(a) Winter (b) Summer

Fig. 3.10: Diurnal variation of GPS-TEC over the Grahamstown station during winter (Figure 3.10(a)) and summer (Figure 3.10(b))

The seasonal variation of the TEC shows that minimum values are observed in winter and summer months with maxima during equinoxes. Figure 3.11 shows an example of the seasonal variation of the TEC over Grahamstown during 2005.

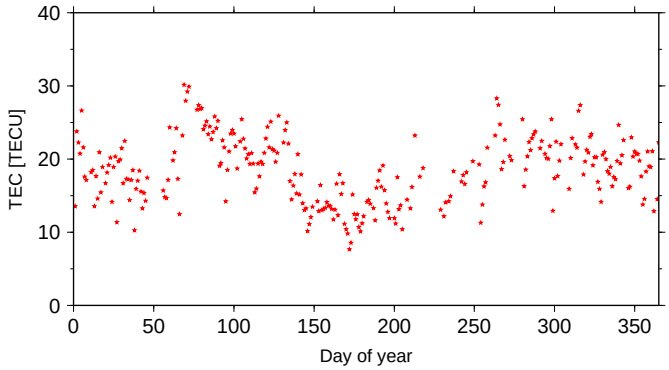


Fig. 3.11: Midday TEC values observed over Grahamstown in 2005. This indicates the seasonal TEC variation over Grahamstown during 2005

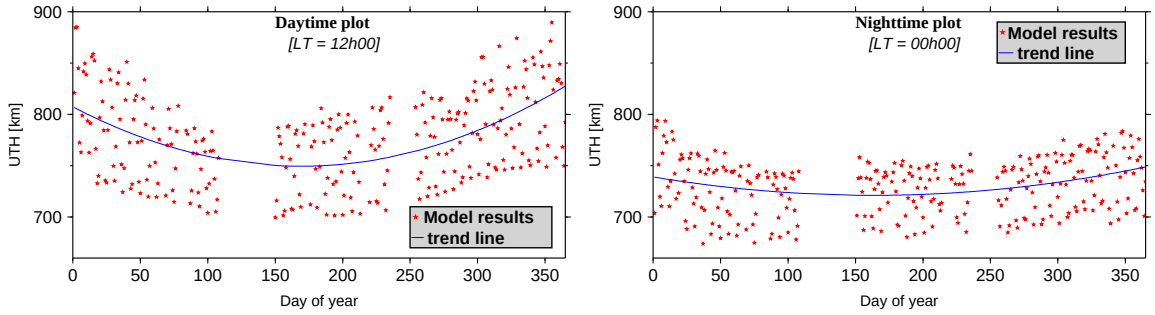
To expand the analysis further, these results are compared with published results of other UTH modelling efforts reported in the literature (Kutiev *et al.* (1994) and Marinov *et al.* (2004)). The published results contain a graphical presentation of generalised variations of the UTH and not incident measurements and reference will be made to the specific publications. In the analysis by Kutiev *et al.* (1994), it was shown that at a station with dip latitude $= 30°$), the UTH follows a diurnal variation with higher UTH values during the daytime and lower values at night. The variation between daytime and nighttime values were different for various seasons with greater differences observed during summer months than during winter months. With respect to solar activity, they observed that the values of the UTH are lower at low solar activity than at high solar activity. Marinov *et al.* (2004) also observed similar UTH behaviour. The magnitudes of the UTH values

are greater during summer than during winter with respect to seasonal variations. Higher values of the UTH occur during the daytime for a mid-latitude region.

This section has highlighted the actual variations of the UTH which are to be considered in more detail in the following sections.

3.6.3.1 Seasonal variation

Figure 3.12 shows examples of the seasonal variation provided by the model at the Grahamstown station (dip-latitude = -50.6°) during a low (Rz12 ≈ 9) solar activity period in 2007. The model values exhibit a minimum during the winter months and a steady increase towards the summer months. This is different from the seasonal trend of the TEC which shows maxima at the equinoxes and lower magnitudes at the winter and summer solstice months as shown in Figure 3.11. The comparison with TEC variation is only for reference purposes and does not in any way indicate error or correctness in the model results. The link between the TEC and UTH is not direct and this study does not attempt to establish that link.



(a) Midday values of UTH; seasonal variation (b) Nighttime seasonal trend of the UTH

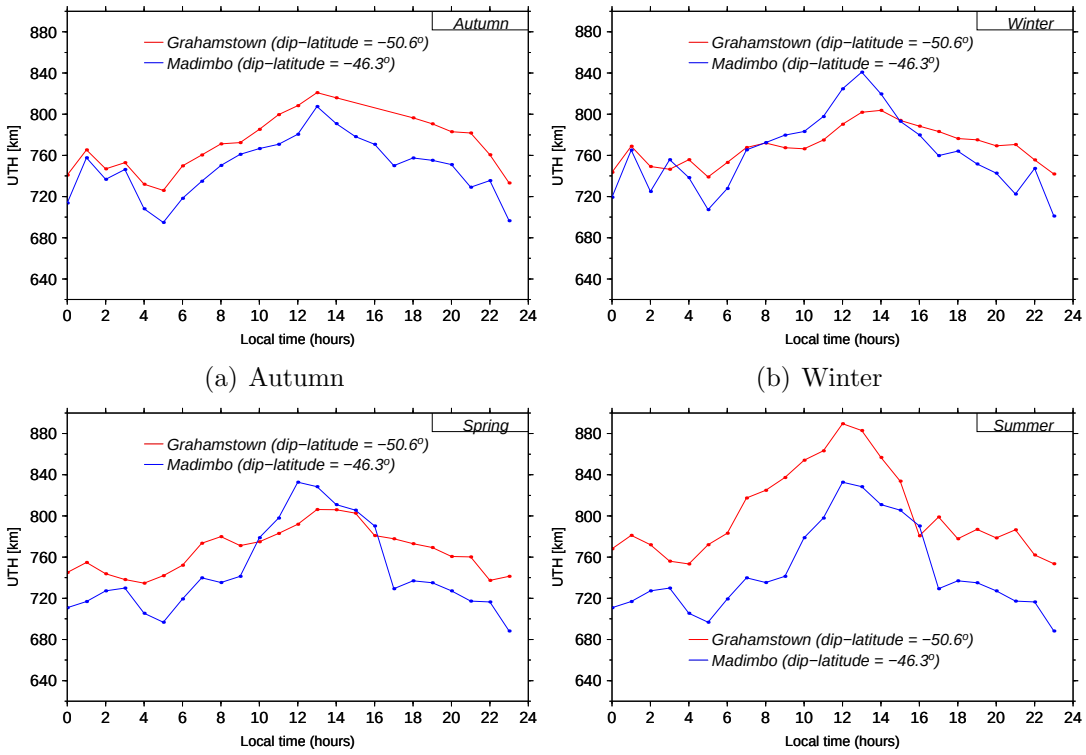
Fig. 3.12: Seasonal variation of the UTH over Grahamstown during a low solar activity period for two local time sectors 00h00 LT and 12h00 LT

The daytime values are marginally higher than the nighttime values. The highest daytime values are observed in summer where the UTH values range between 800 and 900 km. In general over all time sectors, the UTH values range between 650 and 1000 km. A trendline through the data points indicates that the model prediction narrows towards the most likely values of between 700 and 800 km during the low solar activity level, which is physically reasonable and consistent with other observations reported in the literature (Kutiev *et al.*, 1994; Marinov *et al.*, 2004).

3.6.3.2 Diurnal variation

Figure 3.13 presents diurnal variations of the UTH provided by the model over Grahamstown (red line) and Madimbo (blue line) for different seasons: Autumn equinox

(March), Winter solstice (June), Spring equinox (September) and Summer solstice (December). These diurnal variations represent a low (Rz12 ≈ 9) solar activity period during 2007.



(a) Autumn (b) Winter

(c) Spring (d) Summer

Fig. 3.13: Diurnal variations of the UTH during a low (Rz12 < 10) solar activity year for Autumn equinox (March), Winter solstice (June), Spring equinox (September) and Summer solstice (December) at two stations over the South African region: Madimbo (dip-latitude = -46.3°) and Grahamstown (dip-latitude = -50.6°).

These plots of UTH as a function of local time show diurnal variations with a pronounced increase around noon and symmetrical decrease towards the night for all seasons. Generally the day time UTH values are larger than the night time values. The interesting feature observed in all the figures is the minimum between 04h00 and 06h00 LT at both Madimbo and Grahamstown stations. Another distinctive observation worth noting in Figure 3.13 is that the UTH values are significantly larger over Grahamstown than over Madimbo, especially in summer and autumn. This may be attributed to the fact that during summer the Sun is overhead in the Southern hemisphere and during autumn equinox, it is overhead on the equator on its way to the northern hemisphere.

3.6.3.3 Solar activity dependence

It is well known that the temporal variations of the Earth's ionosphere are ultimately coupled to those of solar activity, since the main source of ionisation is the solar

extreme ultraviolet (EUV) and X-ray radiations through the photoionisation of neutrals. High solar activity leads to the enhancement of solar EUV flux, which increases the photoionisation production rate thus contributing to the increase of total ion concentration (or Ne) in the ionosphere. It is expected that the UTH is also closely related to solar activity similar to the F2-peak which moves to higher altitude during higher solar activity.

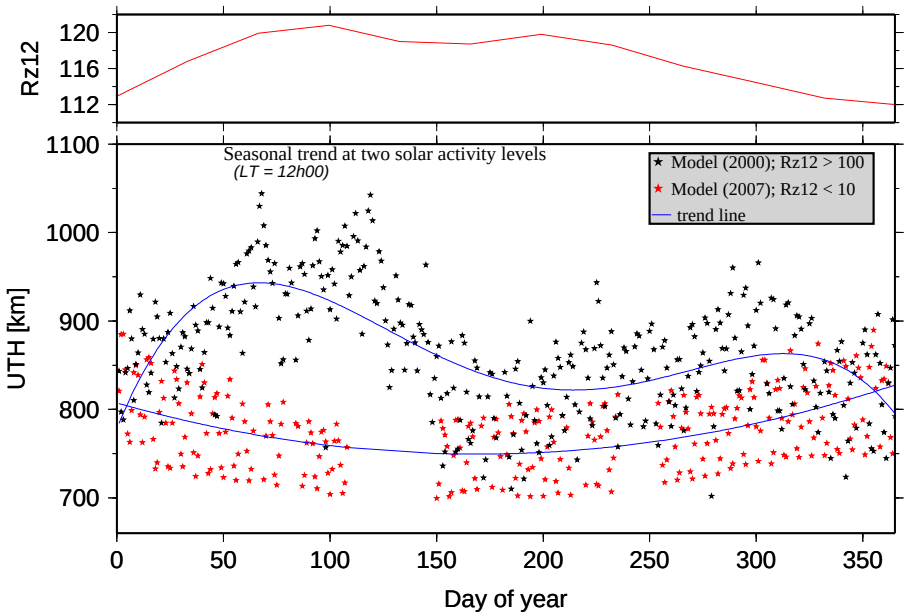
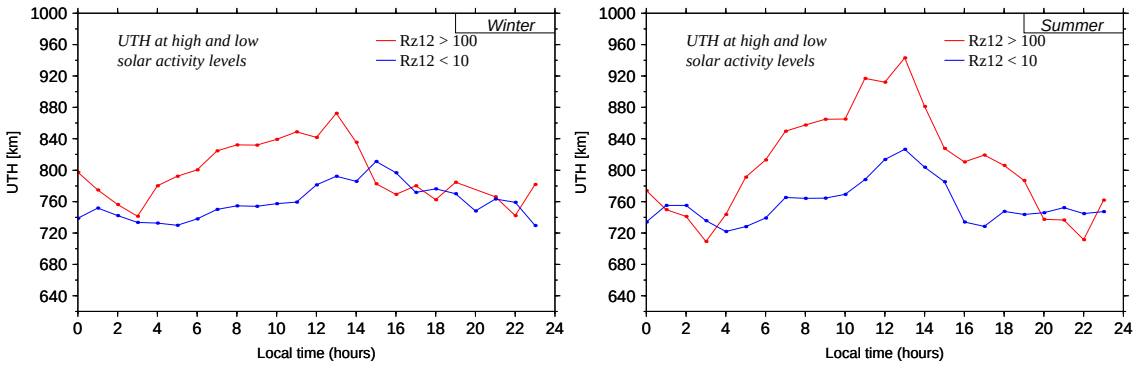


Fig. 3.14: Comparison of seasonal variation between high and low solar activity periods

Figure 3.14 illustrates the seasonal behaviour of the UTH at two levels of solar activity, Rz12 $>$ 100 during 2000 and Rz12 $<$ 10 during 2007, at the Grahamstown ionosonde station. It is clear from the figure that the UTH values are considerably higher during high solar activity than during low solar activity. During the period of low solar activity, the values range between 700 km in winter and about 850 km in summer. During high solar activity, on the other hand, the UTH values range from a minimum of about 800 km to more than 1000 km.

The model predictions with values in the range 650 - 1000 km, although slightly lower, agree in general with results presented by Marinov *et al.* (2004). A significant point to note on Figure 3.14 is the difference in the trend of seasonal variation at the two solar activity levels. Solar activity not only changes the amplitudes, but also drastically changes the way the UTH varies during different seasons. At low solar activity, the seasonal trend shows a minimum in winter and increases steadily towards the summer months, while at high solar activity it exhibits maxima during the equinox months and minima during the solstice months. The reason for this difference is not presently clear.

However, also shown on the same figure (top panel) is a plot of Rz12 versus day of year for the high solar activity period. Rz12 was highest during the Autumn equinox, which corresponds to the largest amplitude in the UTH, and this could be the explanation for different manifestations of the seasonal variations at the two solar activity levels. As Rz12 varied from 112 to 120, there was a corresponding increase in the UTH values of about 200 km.



(a) Winter (b) Summer

Fig. 3.15: Diurnal variations of the UTH at low (Rz12 $<$ 10) and high (Rz12 $>$ 100) solar activity for winter (June) and summer (December) months over Grahamstown (dip-latitude = 50.6°S).

Figure 3.15 shows examples of UTH as a function of local time over Grahamstown for the winter (Figure 3.15(a)) and summer (Figure 3.15(b)) months during high and low solar activity periods. The mean Rz12 values are indicated on each plot. The daytime UTH values during the high solar activity period are larger than those during low solar activity. However, during the night time there is no significant difference in the amplitudes of UTH values between low and high solar activity values.

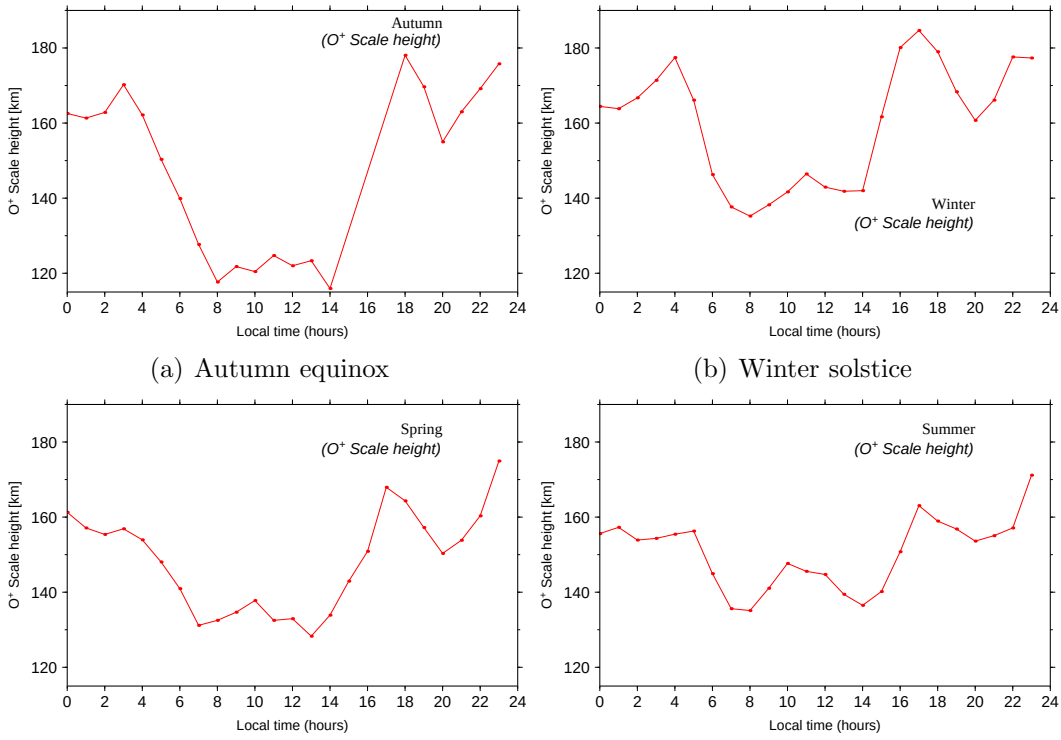
3.6.4 Ion scale heights

There were no measured values available in this study for the ion scale heights to compare with the model results. Similar to the UTH, reference is made to published results of other modelling efforts in order to highlight the differences with the results from the model presented in this thesis. Kutiev *et al.* (2006) approximated the oxygen ion vertical scale height from the topside sounder data, and used this to create a model and provided a comprehensive analysis of their model results. Reference to their work is made in the analysis of the oxygen ion vertical scale height from this neural network based model highlighting the differences between this new model and their models. In addition, reference is also made to Liu *et al.* (2008) to compare their findings with the results from the new model.

3.6.4.1 Diurnal variation

Oxygen ion vertical scale height (H_{O^+})

Figure 3.16 shows the local time variation of the O^+ vertical scale height (H_{O^+}) predicted by the model at Grahamstown during the equinox months (autumn and spring) and the solstice months (winter and summer), representing the different seasons.



(a) Autumn equinox (b) Winter solstice

(c) Spring equinox (d) Summer solstice

Fig. 3.16: Local time variation of the oxygen scale height predicted by the model over Grahamstown, for Autumn (22^{nd} March), Winter (21^{st} June), Spring (20^{th} September) and Summer (20^{th} December), during a low solar activity period (Rz12 $<$ 10).

The local time variation exhibits higher values during the night and a sharp drop to a minimum in the early morning around 06h00 - 08h00 LT thereafter increasing sharply in the afternoon to higher values during the early night time hours. A strange feature is observed during the day where, after the minimum in the early morning, a small and gradual increase occurs post-sunrise to a faint peak around noon and then reduces gradually to another minimum in the early afternoon at around 14h00 - 15h00 LT. After the minimum in the afternoon, the scale height values then increase rapidly and show a pronounced peak around sunset hours (18h00 LT). During the night hours, the scale height again starts to reduce gradually until the pre-midnight hours. This local time variation of the H_{O^+} is evident for all seasons and is more pronounced during the summer.

At this stage the theoretical explanation for this behaviour has not been determined. Theoretically, the plasma scale height (equation 2.5) depends on plasma temperature T_p with a positive correlation. In particular the diurnal variation of the T_p in the upper F-region exhibits a sharp increase after sunrise, due to the heating from fresh photoelectrons and low cooling rate (Brace and Theis, 1981), a phenomenon that is usually called "morning overshoot" (Oyama *et al.*, 1996). The small and gradual increase in the scale height observed in the morning between 09h00 - 12h00 LT may be attributed to the increase in the plasma scale height and/or temperature.

Hydrogen ion vertical scale height (H_{H^+})

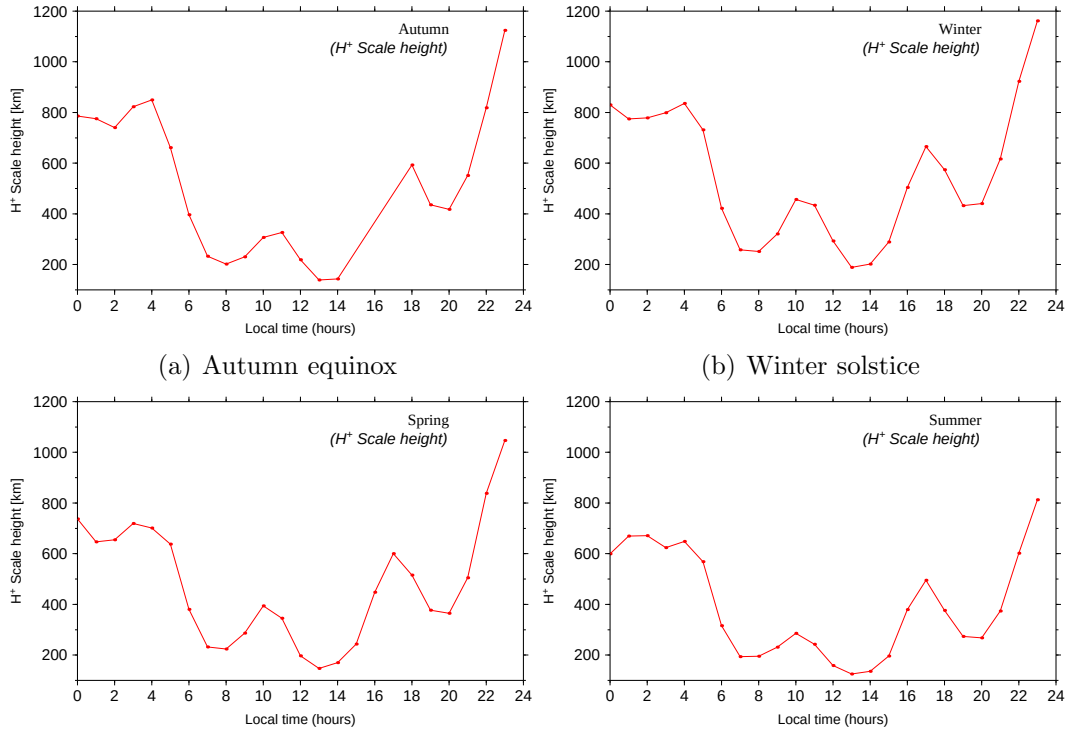
The H_{H^+}, shown in Figure 3.17, also exhibits a diurnal variation similar to that of H_{O^+} and is similar for all the seasons.

In the definition of the scale heights adopted in this work, both the H_{H^+} and the H_{O^+} are related to the gradient of the measured Ne profiles as the vertical distance over which the topside Ne changes by a factor of $e(\approx 2.718282)$, referred to as Vertical Scale Height (VSH), and expressed mathematically as:

$$VSH = -Ne\frac{dh}{d(ln(Ne))}$$

The main difference between the H^+ and O^+ VSH values is that H_{H^+} is the VSH at the upper part of the Ne profile while the H_{O^+} is the VSH at the lower part of the topside ionosphere. Thus, this analysis also reveals the change of the VSH from the value around the F2-peak to the values in the upper topside ionosphere. Theoretically, the scale height is defined as a range of altitudes over which the density along magnetic flux tubes, in the absence of diffusion, decreases by a factor of *"e"*. The scale height is generally defined along the magnetic field lines. The VSH can be obtained from the scale height along the field lines (assuming spherical symmetry) by multiplying it by the cosine of magnetic inclination for the specified location. In this thesis the vertical density gradient, which is derived from the measured Ne profiles, is converted to a scale height dimension.

Kutiev *et al.* (2006) also used measured vertical Ne profiles to derive the vertical scale heights near the F2-peak and developed a model for the VSH. Their method also shows similar characteristics of the VSH over a mid-latitude region (dip latitude = 60°) with lower values during the day and an increase at night. In their results however, there is a minimum during the day around local noon which is quite different from the results



(a) Autumn equinox (b) Winter solstice

(c) Spring equinox (d) Summer solstice

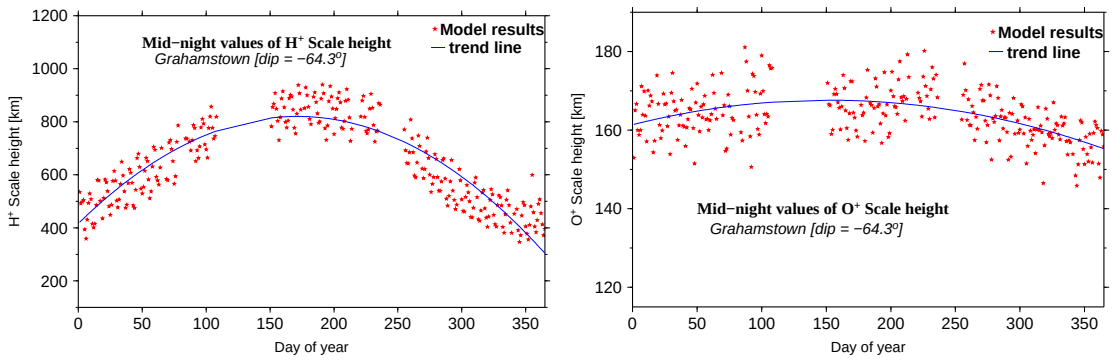
Fig. 3.17: Local time variation of the model hydrogen scale height over Grahamstown, for autumn (22^{nd} March), winter (21^{st} June), spring (20^{th} September) and summer (20^{th} December), with Rz12 < 10.

presented here in which the observations included a minimum in the morning then a gradual increase to a faint peak around noon and another minimum in the afternoon. Liu *et al.* (2008) extracted the vertical scale heights from the vertical Ne profiles of the COSMIC radio occultation measurements during a low solar activity period (2006-2008). They used these values to study the diurnal and seasonal behaviour of the VSH at equatorial and low latitude regions. They did not show any results for mid-latitudes. The morning behaviour of their scale heights over a low latitude region (19°S - 25°S) are comparable to the observation in this thesis which exhibits high VSH values at pre-sunrise hours around 04h00 LT. Then the VSH decreases and approaches a minimum around 07h00 LT. However, their results do not show the kind of features observed in this study during the day, in the afternoon and evening hours. They focused their analysis on the low and equatorial latitudes. The dynamics of the low latitude ionosphere are affected by equatorial phenomena such as the $E \times B$ drift effect which is not as significant at mid-latitudes. The mechanisms that cause the features observed in the diurnal variations of scale heights over Grahamstown, (a mid-latitude region), are unexplained at this time. The nature of the data used may also be a factor in the way the ANN learns to generalise the behavioural pattern of the data it was presented with. As

noted above, the database used does not sample all the longitude sectors uniformly, with more than half of the data obtained for the American sector within a narrow longitude range of 50°W - 104°W. In particular, coverage of the South African region is limited, and due to this the sampling in terms of other parameters such as local time and day of year is fragmented, a fact which may compromise the learning accuracy of the ANN.

3.6.4.2 Seasonal variation

To analyse the seasonal behaviour of the H_{O^+} and H_{H^+} over Grahamstown during low solar activity period (2007), respective plots of the scale heights as a function of day of year for a specific local time sector (midnight) are shown in Figure 3.18. Both the H_{O^+} and the H_{H^+} values are higher during the winter months and lower during the summer months.



(a) Nighttime H_{H^+} seasonal variation (b) Nighttime H_{O^+} seasonal variation

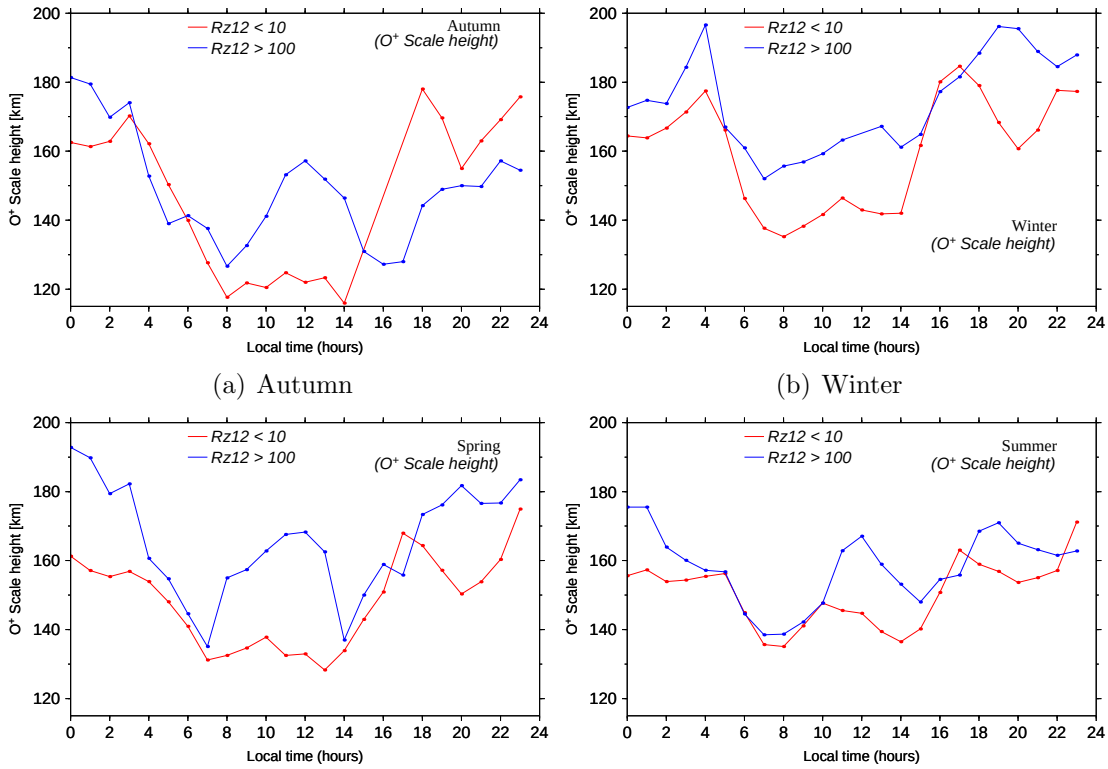
Fig. 3.18: Midnight values of the hydrogen (a) and oxygen (b) ions vertical scale heights as a function of day of year during a low solar activity period.

This was also evident in Figures 3.17 and 3.16 which showed that scale height values are higher during the winter months and lower during the summer. The general diurnal trend, however, is consistent for all seasons.

3.6.4.3 Solar activity effect

In Figure 3.19, the scale height values during the high solar activity periods (blue curves) exhibit similar local time variations with a minimum around 07h00 - 08h00 LT followed by a gradual increase after sunrise to a faint peak around the local noon. Similarly also, the minimum in the early afternoon followed by a sharp rise to a maximum during the evening hours is present during high solar activity period. However, the secondary maximum during the post-sunset hours becomes weak or less significant during the high solar activity period.

Further, it is clear that the vertical scale height values are generally larger during high



(a) Autumn (b) Winter

(c) Spring (d) Summer

Fig. 3.19: Oxygen vertical scale height as a function of local time during different seasons at two levels solar activity

solar activity than during low solar activity for all seasons. This can be understood from the classical definition of the plasma scale height (equation 2.5), in which the scale height is positively correlated with the plasma temperature, which in turn exhibits a linear dependence on the solar activity.

3.7 Procedure to reconstruct the Ne profile

Once the upper transition height, the O^+ and H^+ vertical scale heights are determined, the vertical Ne distribution in the entire altitude range can be constructed. The Ne is given by the sum of the individual ion densities as:

$$Ne(h) = N_{O^+}(h) + N_{H^+}(h) \tag{3.1}$$

where the individual ion densities N_{O^+} and N_{H^+} are assumed to follow an exponential variation. Thus, the Ne as a function of height (h) is given by:

$$Ne(h) = N_{O^+}(h_mF2)exp\left(-\frac{h-h_mF2}{2H_{O^+}}\right) + N_{H^+}(h_mF2)exp\left(-\frac{h-h_mF2}{2H_{H^+}}\right) \tag{3.2}$$

H_{O^+} and H_{H^+} are the oxygen and hydrogen ion vertical scale heights respectively, and $N_{O^+}(h_mF2)$ and $N_{H^+}(h_mF2)$ are the respective densities of the O^+ and H^+ at the height of the F2-peak, h_mF2.

If the ion scale heights are known (from the created model), then the only unknown parameters are the oxygen ion density ($N_{O^+}(h_mF2)$) and the hydrogen ion density ($N_{H^+}(h_mF2)$) at the F2-peak. To determine the unknown parameters, the following principles can be applied:

1.) The principle of plasma quasi-neutrality at h_mF2. The measured Ne is equal to the sum of the individual ion densities. such that:

$$Ne(h_mF2) = N_{O^+}(h_mF2) + N_{H^+}(h_mF2) \tag{3.3}$$

2.) The condition that at the UTH the oxygen and hydrogen ion densities are equal.

$$N_{O^+}(h_mF2)exp\left(-\frac{UTH-h_mF2}{2H_{O^+}}\right) = N_{H^+}(h_mF2)exp\left(-\frac{UTH-h_mF2}{2H_{H^+}}\right) \tag{3.4}$$

Equations 3.3 and 3.4 provide all the information needed to determine the unknown parameters by solving the equations simultaneously and then using them in equation 3.2 to realise the Ne as a function of altitude.

Figure 3.20 illustrates the reconstructed Ne profile compared with the measured profile for the model values of the UTH, H_{O^+} and H_{H^+} which are indicated on the plot.

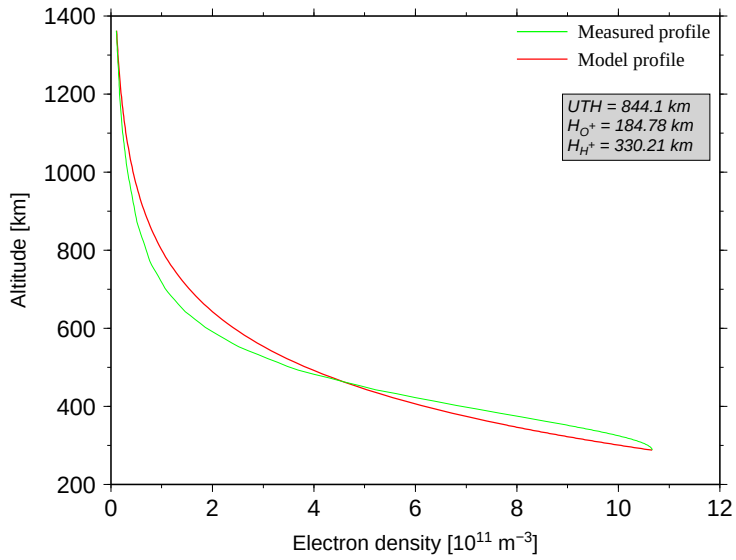


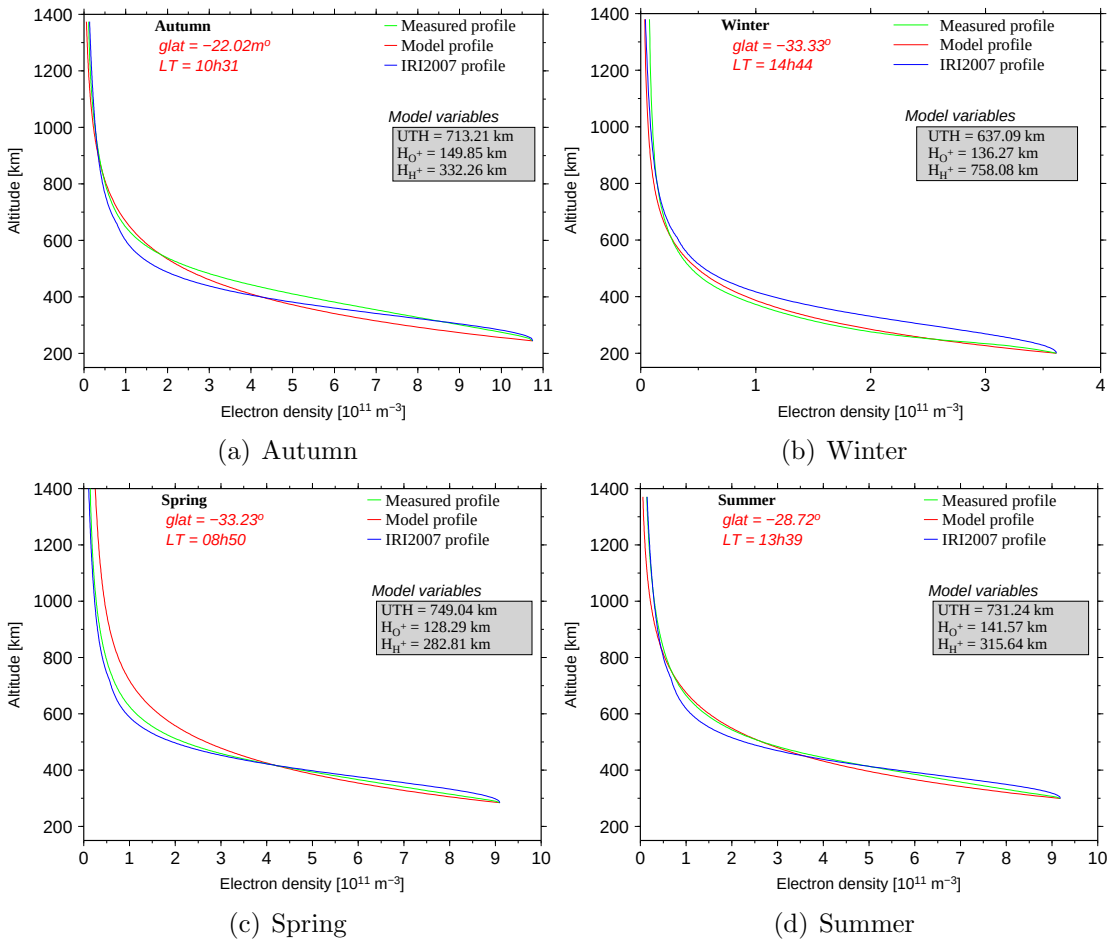
Fig. 3.20: Sample Ne profile reconstruction. The H_{O^+} and H_{H^+} values are indicated with the value of the UTH

3.7.1 Examples of reconstructed topside Ne profiles

The model values of the oxygen scale height, upper transition height and the hydrogen scale height can be used in reconstructing the topside ionosphere electron density profile following the procedure described above. The described reconstruction procedure has been tested on actual measurements of sample topside sounder Ne profiles that fell within the latitude range of the South African region. Figure 3.21 shows the sample reconstructed Ne profiles compared with measured profiles and also with the profiles predicted by the IRI-2007 model.

The sample profiles were chosen to represent each of the four different seasons by a day, and individual profiles from those days corresponding to the daytime sector (08:00 to 14:00 LT) were selected. In addition, the sample profiles were selected to fall within the latitude range of around 20.0° - 36.9°S.

The three model parameters are able to determine the shape of the topside Ne profile which compares well with the actual measurements. The model does not show any systematic discrepancies except some random differences with the measurements such as on Figure 3.21(c) where the model overestimates the measurements around the transition region. It is important to note that the sample profiles fall within the data set which was set aside for testing network and was not used in the training the ANN.



(a) Autumn (b) Winter

(c) Spring (d) Summer

Fig. 3.21: The reconstructed topside electron density profiles (red lines) compared with the measured topside profiles (green lines) and the IRI-2007 model topside profiles (blue lines). The model variables for each plot are indicated on the respective plots. The season, local time and location of the measurements are also indicated on the plots

In essence, the reconstruction procedure will require ionosonde measurements to provide the peak parameters (h_mF2 and $NmF2$) directly, which will ensure the connection of the topside model with the bottomside profile. The topside Ne is considered as a sum of the constituent oxygen and hydrogen ion densities with unknown vertical scale heights.

3.8 Summary

The empirical modelling approach has been used widely and is probably the most practical for ionospheric applications. However, since the approach relies on the availability of good quality measured data, the ionospheric variations may be misrepresented where data is not available. The Alouette and ISIS topside sounder data sets are not uniformly sampled over the essential geophysical conditions, with the largest amount of data originating from the American longitude zone. The South

African region only constitutes about 4.5% of all the available data. Over the years efforts have been devoted to finding ways of using these data sets to develop a global empirical representation of the topside ionosphere. Many of the topside ionospheric models developed using the topside sounder data set were constructed using functional equations. However, in order to construct the functions, the underlying relationship between the independent variables and the dependent variables must be known. This is not always an easy task and in many non-linear situations such as the ionosphere, this may be impossible. One may have to resort to less accurate representations if constrained to write relationships as equations. Neural Networks have been deployed in ionospheric studies (bottomside ionospheric characterisation) to capture the underlying structural relationships between variables, and they have succeeded in providing reliable predictions of the ionospheric behaviour. The main advantage of the neural networks is that no previous knowledge of the nature of the non-linear relationships is required.

In this chapter, strategies of ANN techniques have been employed and their capabilities to provide a representation of the topside ionospheric Ne distribution, based on the Alouette and ISIS topside sounder data, were explored. Three variables, the upper transition height (UTH), oxygen vertical scale height (H_{O^+}) and hydrogen vertical scale height (H_{H^+}), were extracted from each Ne profile. An ANN was trained to construct a representation of these variables as a function of day of year, local time, geographic latitude, geographic longitude, magnetic inclination, 12-month running mean (Rz12) of the sunspot number, height of the F2-peak (h_mF2) and peak Ne ($NmF2$). The values of the UTH, H_{O^+} and H_{H^+} can be used to reconstruct the topside Ne profile using the exponential 'profiler' function. The model is able to provide a fair spatial structure and time variation of the three variables, although the features observed in the model created are quite different from the observations of Kutiev *et al.* (2006) and those of Liu *et al.* (2008). However, the combination of the model variables provided reconstructed electron density profiles that compare well with measurements.

The application of the ANN technique in the topside ionosphere Ne characterisation has challenges. The greatest challenge is the fragmentation of the available data. As noted earlier, the available data sets are limited and sparse. Since ANNs use the data to extract any pre-existing relationships between the input and output variables, when trained on statistically insufficient data, they may not achieve satisfactory generalisation of the relationships. Therefore, although this technique is promising, it is not possible to use this technique optimally without a suitable and more representative database.

Presented in this chapter is a work in progress exploring the capabilities of the ANN

techniques to reconstruct the ionospheric electron density profile based on fragmented topside sounder data. It is important to emphasise that this is the best that can be done with the current available data and the ANN technique. The model variables, especially the UTH, can be useful in other topside ionospheric electron density characterisation approaches, such as using GPS-TEC and other basic measurements to obtain the topside Ne profile as discussed in detail in chapter 5. Presented here are the preliminary results, and the next step will include a more detailed systematic analysis of the model to determine its accuracy, reliability, and suitability for operational application. It is also expected that in time this model will evolve and improve by including other topside sounder data sets from the Intercosmos-19 and the Cosmos-1809 satellites. In addition, if efforts to seek ways of updating the model with other measurements such as the widely available GPS-TEC are successful, it will allow great improvements to the model.

Chapter 4

Theoretical approach

4.1 Introduction

The previous chapter presented an empirical approach which demonstrated the capabilities of neural network techniques to model the topside ionosphere based on topside sounder data, and also highlighted the challenges of using an empirical approach to model an area that suffers from a historical paucity of measured data. Most topside ionospheric modelling studies focused on the empirical approach and it is probably the most practical for ionospheric applications, owing to the fact that the models created are based on measurements which make them more realistic in providing predictions, especially in those areas sufficiently covered by observations.

However, empirical modelling has its own disadvantages. Empirical models in general describe average conditions of the ionosphere since they only attempt to extract the systematic ionospheric variation from past data records. Thus, unless they are adjusted with real-time measurements, they don't provide an accurate specification of the instantaneous ionosphere. Furthermore, since they rely on the availability of good quality measured data, the ionospheric variations may be misrepresented where data is not available. Another limitation is that most available topside sounder measurements only measure up to about 1400 km with very few measurements from the Alouette-2 and ISIS-1 satellites measuring above this altitude. Thus, the empirical models which are based on these data sets can only reliably produce vertical Ne profiles up to an altitude of about 2000 km. As noted, very few measurements of the region above 1400 km have been made. The few known missions include NASA's Radio Plasma Imager (RPI) instrument on the IMAGE satellite that has probed the plasmasphere up to the altitude of 45,922 km which is the satellite's apogee altitude.

The alternative is to develop a model that is based on the mechanisms which control the ionospheric behaviour by modelling the physical and chemical processes that are relevant to this domain of the ionosphere and the plasmasphere. Theoretical models have the potential to be an option for topside modelling and can provide further understanding

of the natural variabilities. The current available theoretical models were not readily available to the author at the time this study was conducted, therefore, analysis to test how such models represent the topside ionosphere in the South African region was not possible. Thus, efforts were made to build the model on first principles and demonstrate the applicability of this approach for the region of interest of this thesis. This chapter presents the attempts undertaken to use theoretical considerations in order to create a model of the altitude variation of the topside ionospheric electron density. This model, based on diffusive equilibrium, is still under development. Therefore the results to be presented are preliminary, and they demonstrate the capabilities of the theoretical approach to construct the topside ionospheric electron density profile.

4.2 General idea to be followed

Theoretical modelling of the topside ionosphere begins with a consideration of the physical processes that are generally accepted to play a significant role in controlling the charged particle distribution in this region of the ionosphere. This approach attempts to solve a set of first principles equations for the ionospheric plasma, which include the continuity, momentum, and energy equations for electrons and ions (Schunk, 1988), as a function of altitude in order to calculate the plasma densities. The greatest challenge with the theoretical approach is that, in order to provide meaningful results, this approach requires a large number of magnetospheric inputs such as electric fields, ion and electron temperatures, neutral temperatures and the oxygen and hydrogen densities, which may not be well known (Cander *et al.*, 1998; Schunk and Sojka, 1992). In addition, the transport equations are coupled, non-linear and time dependent (Anderson, 1993), and therefore the solutions to these equations are not straight forward. Often many simplifications and approximations are made when solving these equations, such as neglecting the presence of certain ion species like He^+, assuming isothermal conditions and approximating the temperature gradients, to name a few.

In view of these limitations, it is important and may be helpful to incorporate measured and/or empirically obtained values of the essential ionospheric parameters into the theoretical considerations to create the model. This work therefore follows an approach that includes a theoretical component with inputs derived from empirical models. Thus the diurnal, seasonal and geographic location variations are contained in the inputs with the effect, this study concentrates on constructing the vertical distribution of the Ne. The model is still under development, and therefore, the results presented are only preliminary, however, these results demonstrate the capabilities of the

diffusive equilibrium approach to provide an option for modelling the topside ionosphere, especially in regions where there are not enough measurements. This chapter also highlights the shortcomings of this approach as a tool to model the topside Ne profile. While this work has not achieved the required results as yet, this chapter provides the starting point for future work to expand the idea presented herein. It also creates an optional working model for the Southern African region where the available measured data is sparse until such time as more measured data becomes available. In the section that follows, a brief discussion on the diffusion theory in the ionosphere is given and the procedure followed in this work is outlined.

4.3 Diffusion theory in the ionosphere

It is a well established concept that diffusion plays an important role in controlling the distribution of ionisation in the ionsphere, and almost completely determines the form of the topside ionosphere (Rishbeth and Garriott, 1969). In the upper region ionosphere or lower plasmasphere where O^+ is still the dominant ion, chemical equilibrium is established controlled by equation 4.1. H^+ is created through this charge exchange with O^+ and the reaction proceeds rapidly in both directions (Webb and Essex, 2000).

$$O^+ + H \rightleftarrows O + H^+ \tag{4.1}$$

This creates ambipolar electric fields between the oxygen ions and the electrons. In the presence of these electric fields, and when the plasma pressure is sufficiently low, the lighter H^+ experiences an outward force and therefore a net outward flow of plasma is established with a density, temperature, and flow velocity consistent with various external influences (Banks *et al.*, 1976). Since the ions are charged, they are constrained to move along geomagnetic field lines. Likewise, when a sufficiently high plasma pressure is present far from the Earth, the H^+ flow can reverse in direction, with H^+ being converted to O^+ within the topside regions above the F2-peak. Finally, a unique state of diffusive equilibrium can prevail such that the gravity and plasma pressure gradients in the O^+, H^+ and other ion gases that may be present precisely balance without flow (Banks *et al.*, 1976).

The earliest known theoretical work on diffusion in the ionosphere was by Hulburt (1928), but at the time there wasn't sufficient knowledge of the ionosphere to accurately assess the importance of diffusion. Ferraro (1945) was the first to show the theory of diffusion as appropriate to the conditions existing in the ionosphere. He considered the ionosphere in diffusive equilibrium as a gas mixture consisting of neutral molecules,

positive ions and electrons, with the electron-ion regarded as a single constituent, and derived the mathematical expressions for diffusive equilibrium in the ionosphere with various simplified conditions (Rishbeth, 1975).

Since then, other studies considered the problem and in the 1950s solutions of the diffusion equations representative of the real ionosphere were obtained for various situations, although they were also idealised. For example, Yonezawa (1955) and Dungey (1956) considered diffusion, loss and vertical electrogeomagnetic drift in the F2-layer during nighttime, and Yonezawa (1958) obtained a solution for the daytime condition assuming a steady-state condition and the presence of vertical electrogeomagnetic drift. Other later studies on the subject proposed modifications to the original theory of Ferraro to take the temperature gradient and the effect of horizontal gradients into account (e.g Bauer (1969)).

Several studies followed the theoretical approach to create models of the electron density as a function of altitude and provided detailed discussions on the theory of diffusion in the ionosphere. This section presents only a brief overview of the diffusive equilibrium theory. For more in-depth study the reader is referred to the cited works, especially Rishbeth and Garriott (1969) and Bauer (1969).

Various works applied the diffusive equilibrium approach differently in efforts to characterise the vertical variation of the ionospheric Ne. For example, Bauer (1962b) derived solutions to the diffusive equilibrium problem for a ternary ion mixture of O^+, H^+ and He^+ and expressed the Ne in terms of the mean ionic mass $\bar{m} = \frac{\sum m_i n_i}{\sum n_i}$ given by:

$$Ne = N_o exp\left(-\frac{g_o}{2k_b T_p}\int_0^{z'} \bar{m} dz'\right) \tag{4.2}$$

where g_o is the acceleration due to gravity at the Earth's surface, k_b is the Boltzmann's constant, T_p is the plasma temperature and z' is the height parameter expressed in terms of the geopotential altitude $h' = \int_0^h g/g_o dh$. n_i is the ion species that contributes to the computation of $\bar{m}$ given by

$$n_i = n_{io} exp\left(1 - \frac{z'}{H_i}\right) exp\left(\frac{g_o}{2k_b T_p}\int_0^{z'} \bar{m} dz'\right) \tag{4.3}$$

and the scale height is $H_i = k_b T_p / m_i g$. Bauer (1962a) later applied this approach to a binary ion mixture and showed the effect of the relative abundance of the ions in controlling the shape of the resulting topside Ne profiles. The main difficulty with this

approach is that the mean ionic mass is not a physical parameter. It is usually expressed in terms of the relative abundance of the individual ions and can be determined by fitting theoretical models to observed density profiles (Bauer, 1969).

Titheridge (1972) also provided solutions to the diffusive equilibrium equations for a mixture of O^+ and H^+. His solutions are discussed in section 4.3.3 where the Ne is expressed in terms of the ion and electron temperatures. He showed that while the electron temperature plays a role in controlling the shape of the topside profile, the relative abundance of the ions plays a major role.

In more recent studies, Webb and Essex (2000), implemented Titheridge's solutions with modified temperature profiles and developed a global model of the Ne along magnetic field lines. For the lighter ions, they used chemical equilibrium at lower altitudes and diffusive equilibrium at higher altitudes, using the equation derived by Richards and Torr (1985) to determine the altitude at which to switch from chemical equilibrium to diffusive equilibrium.

Yoshimura *et al.* (2005) combined theoretical considerations with GPS and ionosonde measurements to create a semi-theoretical representation of the plasma distribution in the topside ionosphere.

Diffusive equilibrium is also successfully implemented in the *Cost Prof* model (Leitinger, 1998; Leitinger *et al.*, 2002) in which the topside Ne formulation is given in terms of (a) the oxygen scale height at the F2-peak, (b) its gradient and (c) the upper transition height. The three parameters are modeled as functions of solar activity, local time, season and modified dip latitude.

4.3.1 Continuity equation

Starting with a simple situation of electrons and ions as constituting a gas (a plasma) of concentration N, which is a minor constituent of the atmosphere, the continuity equation for the diffusion of such a minor gas through a predominant background gas is given by

$$\frac{\partial N}{\partial t} = q - \bigtriangledown(N\nu) - l(N) \tag{4.4}$$

Where $\bigtriangledown(N\nu)$ represents the source or sink as the result of mass transport, where ν is the transport velocity, t is the time, q is the rate of ion pair production and l (a function of N) is the loss rate of charged particles due to chemical processes. The equation contains the derivatives with respect to time and space. In the case of a large structure such as

the ionosphere, the equation is more complicated since the plasma is subject to various conditions such as gravity which introduces the scale height (Rishbeth, 1975), collisional forces which introduce collisional terms, electric forces between the ions and the much lighter and more mobile electrons, the geomagnetic forces that constrain the diffusion along geomagnetic field lines, reducing the coefficient of diffusion by a factor of $\sin^2 I$, I the inclination (dip) of the Earth's geomagnetic field. In addition, the decrease of air density with increasing altitude results in a decrease in the diffusion coefficient. These factors introduce extra terms into the equation which will not be discussed in this thesis. Readers are referred to Rishbeth (1975) for more details. In Bauer (1969) it is shown that if the horizontal motions are neglected, since they contribute very little compared to the the vertical motions, the condition for diffusive equilibrium can be expressed by:

$$\bigtriangledown(N\nu) \cong \frac{\partial N\nu_z}{\partial z} = 0 \tag{4.5}$$

where ν_z is the upward drift velocity and z is the height. Bauer (1969) showed further that the solution: $\nu = 0$ for equation 4.5 satisfies a situation where the plasma reached a stable distribution as a result of diffusion.

4.3.2 The basic diffusion equations

The derivation of the basic equations for the diffusion of ions and electrons in the ionosphere is given by Rishbeth and Garriott (1969), and the equations are reproduced here as equations 4.6 and 4.7 for the ions and the electrons respectively. The partial pressures are balanced by forces due to gravity g (first terms on the right hand side), and an electrostatic polarisation field E, the collisional forces between the ions and neutrals with collisional frequency ω_{in}; and the collisional forces between electrons and neutrals with collisional frequency ω_{en} so that the diffusion equations for the vertical motions can be expressed by:

$$\frac{d(N_i k_b T_i)}{dh} = -N_i m_i g + N_i eE - N_i m_i \omega_{in}(\nu_i - \nu_n) \tag{4.6}$$

$$\frac{d(N_e k_b T_e)}{dh} = -N_e m_e g - N_i eE - N_e m_e \omega_{en}(\nu_e - \nu_n), \tag{4.7}$$

Here k_b is the Boltzmann constant, N is the number density, T the temperature of the appropriate charged constituent according to the subscript (e for electrons and i for ions), e is the charge of the electrons and ν denotes vertical drift velocity (with subscripts i, e, n for ions, electrons and neutrals) respectively. These equations serve as a starting

point for discussing the equilibrium distribution of the charged particles in the topside ionosphere.

4.3.3 Diffusive equilibrium

From section 4.3.2, it is shown that diffusive equilibrium in a plasma in the topside ionosphere can be expressed by force equations for the charged particles. Upon neglecting all terms containing collisional frequencies, since at these heights the collisions are negligible, the diffusion equations reduce to partial pressure, gravitational and electric field terms. Furthermore also, neglecting the electron mass which is negligibly small in comparison with the ion mass m_i, and taking into account the fact that several positive ion species may be present, the equations 4.6 and 4.7 for the j^{th} ion and for the electrons can be re-written as:

$$\frac{d(N_j k_b T_i)}{dh} = \bigtriangledown(p_i) = -N_j m_j g - N_j eE \tag{4.8}$$

$$\frac{d(N k_b T_p)}{dh} = \bigtriangledown(p_e) = -NeE \tag{4.9}$$

where p_i and p_e are the respective pressures for the ions and electrons, k_b is the Boltzmann constant, N_j is number density for the j^{th} ions and N is number density for the electrons and T (e for electrons and i for ions) are the temperatures of the appropriately charged particles. In the section that follows, these equations are used to create a model.

4.4 Procedure followed to create the model

The problem, as defined earlier, is to determine the topside ionosphere electron density distribution using theoretical considerations of physical and chemical processes which are responsible for the distribution. A model of the topside ionosphere will be developed based on diffusive equilibrium. In general diffusive equilibrium should provide a good description of the topside ionosphere. However, from various observations and theoretical considerations reported in the literature (Bauer, 1969; Webb and Essex, 2000), diffusive equilibrium for the lighter ions (H^+) cannot be assumed to be the case at lower altitudes. Therefore, chemical equilibrium is used to model the H^+ in the lower part of the topside ionosphere and diffusive equilibrium is used for the upper part.

4.4.1 Description of assumptions and approximations

To achieve the intended goal and draw meaningful conclusions from the approach, the assumptions listed below will be needed throughout the study. Comments on the justification for each assumption are given based on arguments derived from a consideration of the physics of the ionosphere and results reported in various literature:

- Diffusion is assumed to be the dominant process controlling the distribution of electrons and ions, i.e each ionic constituent itself follows a diffusive equilibrium distribution. As pointed out earlier, plasma distributions in the topside ionosphere are controlled by the plasma transport processes where field-aligned upward plasma flows supply plasma in the plasmasphere during the daytime and the downward flows contribute to maintain the nighttime F-region.

- It is assumed that the topside ionosphere consists only of O^+ and H^+ ion species. The presence of He^+ and other ions are neglected. Based on observations by Carlson and Gordon (1966) and Hoffman (1967), they have little effect on the electron density profiles. It is further assumed that the ionosphere is statistically neutral i.e., the ions are singly charged so that the sum of the ion densities equals the electron density.

- It is also assumed that the ionosphere is horizontally stratified, i.e. horizontal gradients of all quantities are much smaller than vertical gradients (Bauer, 1969).

4.5 Solutions to the diffusive equilibrium equations

Solutions to the diffusive equilibrium equations 4.8 and 4.9 derived by Titheridge (1972), as reproduced below, were adopted for this work. The solutions provide a definitive expression of the scale height for the j^{th} ion species with mass m_j in terms of electron temperature (T_e), ion temperature (T_i), and mean ionic mass ($\bar{m}$) given by:

$$H_j = -\frac{k_b T_i}{m_j g}\left(1 - \frac{\bar{m} T_e}{m_j T_t}\right)^{-1} \tag{4.10}$$

where, k_b is the Boltzmann constant, g is the gravitational acceleration and $T_t = T_i + T_e$. The electric field E was eliminated by considering:

- charge equilibrium and summing over all ions, $\sum_j N_j = N_e$,

- the presence of more than one ionic constituent which gave rise to the mean ionic

mass, given by:

$$\bar{m} = \frac{\sum_j N_j m_j}{\sum_j N_j} \tag{4.11}$$

Since the mean ionic mass is not a "physical" parameter, it is often of interest to represent this parameter in the form of ion composition, i.e., the relative abundance of individual ions.

The solution also takes the existence of temperature gradients into account. Hence, temperatures at each altitude level must be known to compute the electron density. It is clear from the scale height expression that in the general case, a physical interpretation of electron density profiles requires knowledge of a number of other parameters, such as ion and electron temperatures and mean ionic mass.

Thus, if the number density N_{jo} of the j^{th} ion species at a base height h_o with temperature T_{to} is known, the density N_j at some greater height $h_o + \triangle h$ can be calculated from

$$N_j T_t = N_{jo} T_{to} exp(- \triangle h / H_j) \tag{4.12}$$

This can be done by numerical integration (Titheridge, 1972; Webb and Essex, 2000), such that, in a series of small altitude "steps", the number density of a given ion species is determined at each altitude level given (i) its density at a certain base height, (ii) the mean ionic mass and (iii) the ion and electron temperatures are specified at each altitude level.

Plasma is constrained to move along geomagnetic field lines since diffusion of the charged particles is governed by the geomagnetic field. Thus, simply projecting up from the ionosphere assuming chemical and diffusive equilibrium will significantly over estimate the correct plasmaspheric density (Webb and Essex, 2000). The diffusion profiles must therefore be calculated along the field lines rather than simply in the vertical direction. In this study, the vertical scale height was mapped to the field aligned scale height at each calculation point via the geometry of the geomagnetic field. Given ds, the differential element along a magnetic field line, and I the magnetic inclination, it follows by simple geometrical considerations that the differential element in the vertical is $dh = \sin I ds$ (Stankov *et al.*, 2003). Using the relationship between the magnetic dip-latitude and the inclination for a dipole, given by $\tan\lambda = 1/2 \tan I$ (Chapman, 1963) where λ is the magnetic dip-latitude, the correction factor is given by:

$$\tau = \sin(\arctan[2\tan(\lambda)]) \tag{4.13}$$

The chemical equilibrium profiles for the H^+ were calculated using the equation;

$$N_{H^+} = \frac{9}{8}(N_H \cdot N_{O^+})/N_O \tag{4.14}$$

where N_H is the number density for atomic hydrogen, N_O is the number density for oxygen atoms at the reference height and N_{O^+} is the oxygen ion density.

4.5.1 Calculating the *Ne* profile

The topside ionospheric electron density profiles were calculated using equation 4.12. This calculation requires three outside sources of input parameters:

1. Electron and ion temperatures (T_e and T_i) at each height
2. The ion number densities (N_{H^+} and N_{O^+}) at an arbitrary base height
3. The oxygen and hydrogen neutral number densities

The required parameters were obtained from empirical models. T_e and T_i were obtained from the IRI model, however, temperature profiles of the IRI model extend only up to 3000 km (Bilitza, 1985). In the IRI model T_e is defined by global fits to independent data sets, at heights of 400, 600, 1400 and 3000 km, with linear interpolation between these heights (Bilitza, 1991a). Titheridge (1998) observed that as a result of this the gradients dT_e/dh are sometimes unrealistic, with large discontinuities. The ion temperature T_i is modelled independently with the range limited by the values of T_e and the neutral temperature T_n, such that $T_i = T_n$ at lower altitudes and becomes equal to T_e in the plasmasphere. Examples of the electron and ion temperature profiles from the IRI model are shown in Figure 4.1.

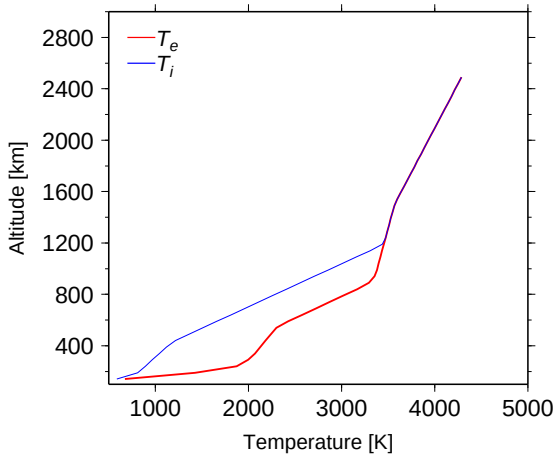


Fig. 4.1: Example of electron and ion temperature profile from the IRI model

Titheridge (1998) developed a model of the electron and ion temperatures based on data from the EXOS-D satellite that has both a theoretical and empirical component and extends up to 8000 km. The model provides an analytic representation of the temperatures as functions of height along the magnetic field lines. The model solves the heat flow equations and then fits these equations to the EXOS-D satellite data by adjusting two parameters of the temperature expression. Access to this model was not possible at the time of preparation of this thesis.

The neutral parameters (N_H and N_O) were obtained from the global atmospheric model, Mass Spectrometer and Incoherent Scatter (MSIS-90), (Hedin, 1991), which provides the atmospheric parameters as a function of latitude, longitude, height and solar activity. The model is based on a large amount of satellite and backscatter data. The $NmF2$ and h_mF2 obtained from topside sounder observations for each profile example were used to provide the base height and the ion densities, N_0 and N_H, at the reference height h_0. The relative abundance of the ions were provided by the IRI model. The starting height for the diffusive equilibrium calculation was taken to be at about one scale height above the peak of the F2-layer, h_mF2. For the region below this altitude, an α-Chapman layer with a scale height $H_c = H_O(h_mF2)$ was used to calculate the Ne to force a peak, since diffusive equilibrium does not give a minimum but a decrease in the electron density with an increase in height.

4.6 Results and discussions

As pointed out earlier, the model is still under development, and therefore the results presented here are preliminary. However, they indicate the type of results that the model is capable of producing, noting that its primary goal is to calculate the electron density as a function of altitude. The most fundamental results required from the model is the vertical distribution of the Ne at a given time at a point specified by latitude and longitude. An example of this is a daytime (12h00 LT) diffusive equilibrium electron density profile for a mid-latitude location (33.3°S, 26.5°E) on 16 November 2005, as shown in Figure 4.2.

From this result it would appear that diffusive equilibrium should provide a general description of the Ne distribution in the topside ionosphere. In deriving the Ne it is also assumed that the individual ion constituents that contribute to the electron density distribution follow a diffusive equilibrium distribution, with the O^+ being the dominant ion in the topside and the H^+ becoming dominant at higher altitudes.

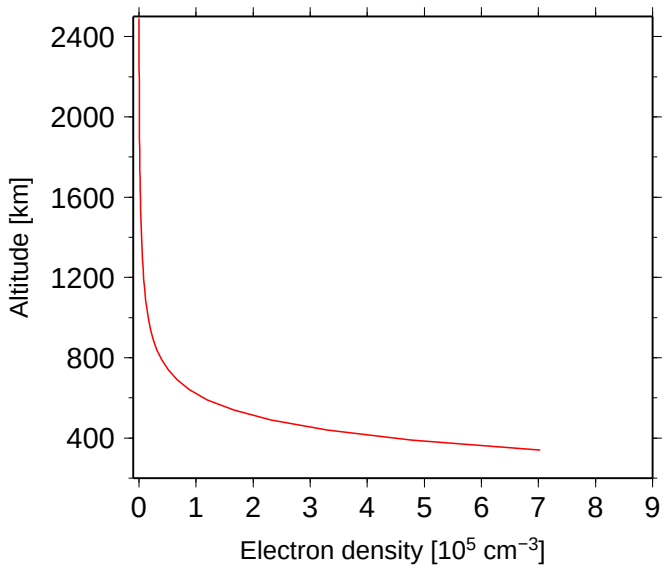


Fig. 4.2: An example of an electron density profile derived with the diffusive equilibrium approach

Shown in Figure 4.3 are the scale height profiles for the O^+ and H^+ corresponding to the electron density profile given in Figure 4.2.

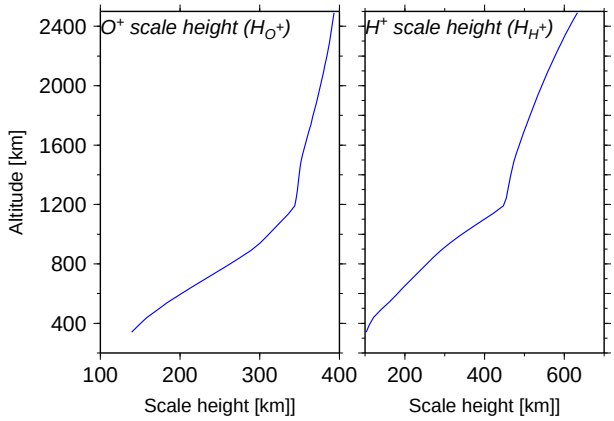
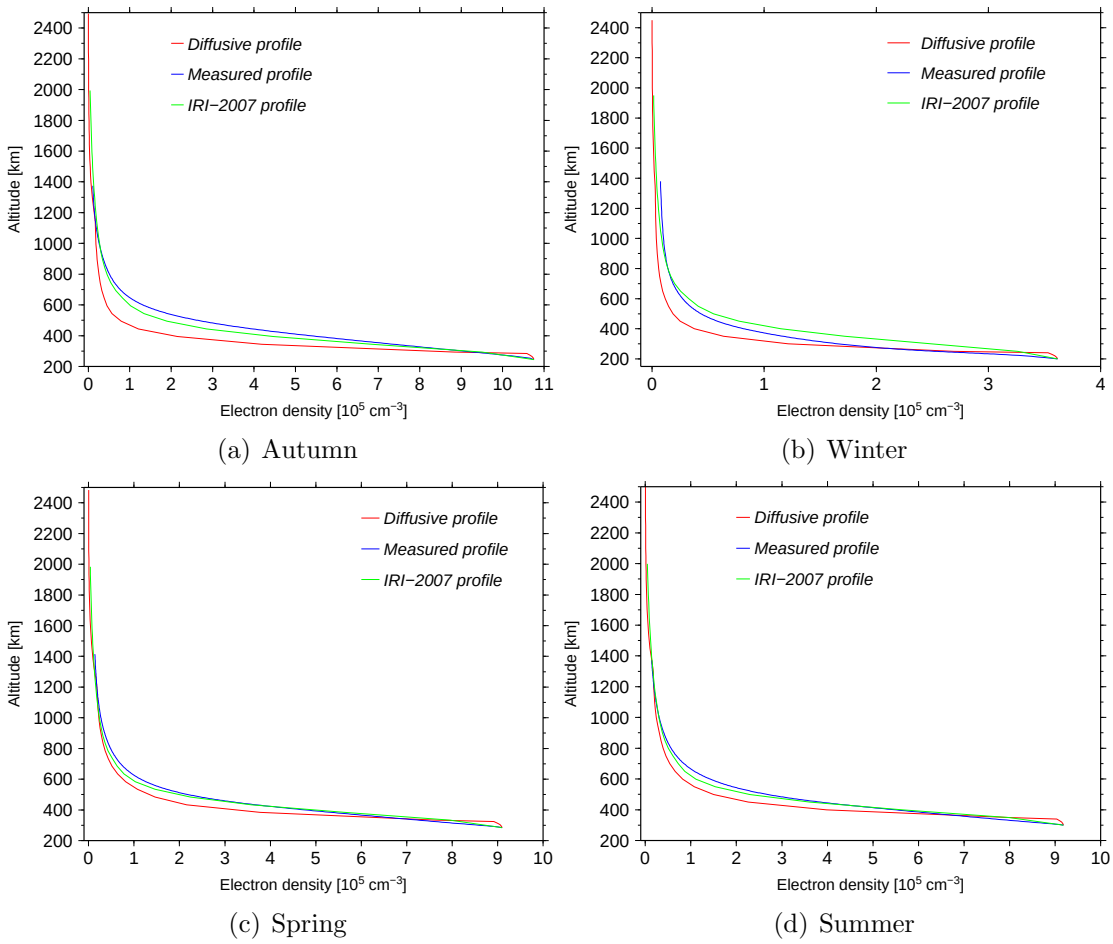


Fig. 4.3: Example of O^+ and H^+ scale heights as a function of altitude calculated using equation 4.10

Figure 4.4 shows a comparison of the diffusive equilibrium profiles with measured profiles. A comparison of the diffusive equilibrium profiles with the empirical profiles from the IRI-2007 model, the most commonly used ionospheric model to predict the topside ionosphere over South Africa is also provided. The profiles shown all fell within the latitude range of around 20.0° - 36.9°S and each of the four seasons is represented by a day, with the individual profiles from those days corresponding to the daytime sector (08:00 to 14:00 LT).

Although the diffusive equilibrium profiles generally follow the shape of the topside compared with the measured profiles, there are significant disagreements between the



(a) Autumn (b) Winter

(c) Spring (d) Summer

Fig. 4.4: Theoretical diffusive equilibrium vertical Ne profiles for autumn (March), winter (June), spring (September) and summer (December) compared with the measured profiles (blue curves) and the IRI profiles (green curves)

diffusive equilibrium profiles and the measured profiles around the transition region, where the profiles change shape with the diffusive equilibrium underestimating the measured profiles. This disagreement may be attributed to the temperature model used. Since the variation of electron density with height is controlled closely by the effects of diffusion, the electron and ion temperatures (T_e and T_i) play a key role in controlling the gradient dN_e/dh. The temperature profiles from the IRI model have discontinuities due to the interpolation method used, and as Titheridge (1998) observed, the gradients dT_e/dh are sometimes unrealistic. In addition, the discrepancy may also be attributed to the ion composition values used at the base height. As Titheridge (1972) noted, the relative abundance of the ions play a major role in controlling the shape of the topside profile.

4.6.1 Discussion

Presented in this chapter is an overview of some of the possible results from a model that is based on diffusive equilibrium theory. The results obtained agree in general with experimental observations (for example Figure 4.4, showing a comparison with the ISIS-2 Ne profiles and the IRI-2007 empirical model). These results demonstrate that the diffusive equilibrium theory can be a useful model concept in modelling the topside ionosphere. Generally, it appears that a diffusive equilibrium distribution represents a reasonable model of the topside ionosphere. However, this needs to be tested on more recent measurements. Thus, even though the model may appear to work on the older profiles that are available there is no guarantee at this stage that the model would be accurate enough for current conditions.

4.7 Summary and limitations

The advantage of the theoretical approach in modelling the ionosphere is that the response of the system to magnetospheric phenomena, such as geomagnetic storms and substorms, can be modelled. The main disadvantage of the theoretical approach is that a large number of inputs, such as ion and electron temperatures, neutral temperature and the relative abundances of the constituent ion densities are required, which may not be well known (Cander *et al.*, 1998). In the case of the South African region, these parameters are not, at this stage, well studied. Generally, these inputs are derived from other models. Thus in addition to the limitations inherent within the approach due to the assumptions and generalisations made, the results will also depend on the accuracy of the models that provide the inputs. Further work should involve exploring ways to incorporate more readily available measurements such as GPS-TEC instead. The diffusive equilibrium concept in general represents an idealised situation physically, and an asymptotic solution mathematically, so that the derived quantities are not necessarily unique, although they may appear to fit the observed profiles (Webb and Essex, 2000).

Chapter 5

Characterising the topside Ne using GPS data

5.1 Introduction

As shown in chapter three, the available measured topside Ne data sets provide a limited coverage of the relevant geophysical conditions (Benson *et al.*, 1998) and the data are not sufficient to properly characterise the structure of the topside ionosphere under different conditions such as altitude, geographical location, as well as diurnal, seasonal, solar activity and geomagnetic activity effects. The ionograms from the Alouette and ISIS satellites that were processed into Ne profiles represent only a small percentage of all the measurements that were made (Huang *et al.*, 2002). More than 50% of the processed Ne profiles fall within a narrow longitude range of 50°W - 104°W, leaving the rest of the longitude zones with less than 50% of the data, and coverage of the South African region is sparse within the database.

The small amount of measured topside data available for the South African region and its distribution over the various geophysical conditions presents a challenge in topside modelling efforts for this region. This emphasised the need to look for other data sources for topside Ne modelling in this region.

Over the past two decades, the Global Navigation Satellite Systems (GNSS), such as the Global Positioning System (GPS), have become a promising tool for retrieving ionospheric features under different conditions. The use of GPS observations as a tool for ionospheric characterisation takes advantage of the dispersive propagation properties of the ionosphere for radio signals. For microwaves, such as the GPS signals travelling from the GPS satellites through the ionosphere to the GPS receiver on the Earth's surface, the ionosphere acts as a dispersive medium (Komjathy, 1997). The impact on the GPS signals is a function of the carrier frequency and the Ne along the signal path. Using this ionospheric effect on the GPS signals, dual-frequency GPS observations can be used to determine the Total Electron Content (TEC), i.e. the integral of the Ne along

the ray-path between the transmitting satellite and the receiver. TEC is an important characteristic of the Earth's ionosphere that carries information on time and position variability of the ionosphere and has proved to be useful as a sensor of ionospheric climatology (Davies and Hartmann, 1997; Jakowski *et al.*, 2004; Rama Rao *et al.*, 2006). Applications include data assimilation techniques in ionospheric modelling where the GPS-TEC can be used as an anchor point to adapt models for the locations and epochs of interest (Heise *et al.*, 2003; Liu *et al.*, 2005).

However, such satellite to ground-based receiver measurements can only produce information about the density in the form of path integrated snap-shots of the TEC. The challenge is to decorrelate this to generate the distribution of the Ne with altitude. In recent years, several studies have been undertaken to develop techniques for inferring the Ne distribution in the ionosphere from GPS-TEC measurements. This chapter provides an overview of the recent progress in topside ionospheric Ne modelling techniques that are based on GPS-TEC and explores the possibility of using GPS measurements for the topside modelling efforts in South Africa.

Recently, the Chief Directorate Surveys and Mapping (CDSM) of South Africa set up a network of dual-frequency continuously operating GPS base stations (Trignet network) distributed throughout South Africa at approximately a 200 - 300 km spacing, as shown in Figure 5.1.

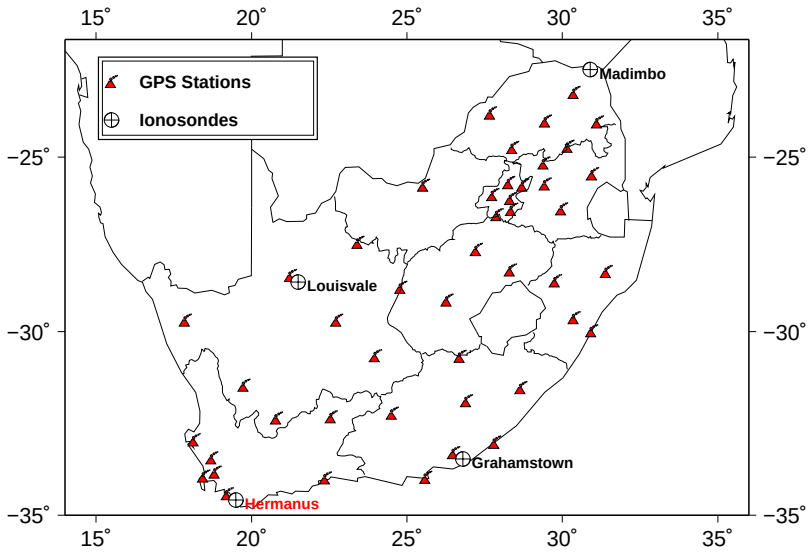


Fig. 5.1: CDSM GPS network (red) triangles and the four ionosondes are indicated. Note the co-located GPS receivers at Grahamstown, Louisvale and Hermanus ionospheric stations.

The data are made available to the scientific community and have presented unprecedented opportunities for ionospheric studies. These GPS-TEC measurements are

a new data resource that can be used with other ionospheric measurements to monitor the actual state of the ionosphere continuously and to characterise the structure of topside ionospheric Ne reliably in this region.

In this chapter, the procedure to reconstruct the topside Ne profile using a combination of GPS-TEC and bottomside ionosonde observations first proposed by Stankov and Muhtarov (2001) is implemented to provide a representation of the topside Ne over the Grahamstown (33.3°S, 26.5°E) ionospheric station. The reconstructed Ne profile is tied to the measured GPS-TEC and the F2-peak parameters (h_mF2, and $NmF2$), as well as the bottomside TEC, with the effect that the variation of the reconstructed profile in terms of local time, season, geographic location, solar activity and geomagnetic activity are contained within the measured parameters. Thus, the study focuses on the altitude distribution of the Ne.

This chapter is organised into four main parts.

- First, an overview of the use of GPS as a tool for ionospheric characterisation is given, highlighting the basics of how TEC is derived from the GPS measurements. In addition, some general variabilities of TEC are given, i.e. the typical diurnal, seasonal and geographical location. Reference is made to various groups that have provided detailed studies of TEC variation.

- Secondly, an overview of the various approaches that have been developed to derive the vertical distribution of the Ne from the GPS-TEC measurements is presented.

- Then a description of the procedure used in this study to reconstruct the topside Ne profiles from a combination of GPS-TEC data, ionosonde measurements and model values of the upper transition height (UTH) is presented.

- Finally, the results obtained and the procedures applied are discussed.

5.2 GPS as a tool for ionospheric characterisation

As pointed out earlier, ionospheric TEC is defined as the number of free electrons contained in a column of unit cross-section area (Suard and Arbesser-Rastburg, 2005), and is usually expressed in TEC units (TECU) where 1 TECU = $10^{16}el/m^3$. In mathematical form, it is expressed as:

$$TEC = \int_p n(s)ds \tag{5.1}$$

where $n(s)$ is the Ne and p is the propagation path between the satellite and the receiver. Since the late 1990s, the GPS system has been used to provide a measure of the TEC along a ray path between the satellite and the receiver. Each GPS satellite transmits information for positioning purposes on two carrier frequencies, $L1$ and $L2$, centred on 1.57542 and 1.22760 GHz respectively (Bassiri and Hajj, 1992; Jeffrey *et al.*, 1988). Due to its dispersive nature, the ionosphere causes path delays on the carrier frequencies which depend on the carrier frequencies themselves (Calais and Minister, 1998) and the Ne along the ray path. Simultaneous observations of the two carrier frequencies with dual-frequency GPS receivers allow for the elimination of this ionospheric effect. It is this effect of the ionosphere on the GPS signals that is useful for ionospheric studies.

The GPS measurements provide the pseudo ranges ($P1$ and $P2$) and the carrier phases ($L1$ and $L2$), which in simple terms provide measurements of the slant range between the GPS satellite and the receiver. The difference between the $P1$ and $P2$ pseudo ranges expressed in units of length allows for information on the integrated Ne along the signal propagation path to be extracted from the dual-frequency GPS observations. The derived TEC has to be calibrated for the receiver and satellite differential group delay biases. Details on how the TEC is calculated from the GPS data, as well as the estimation of the ground receiver and satellite biases, are given in Opperman (2007). The section that follows provides a brief overview of the methodology followed by Opperman (2007).

5.2.1 Adjusted Spherical Harmonic Analysis (ASHA) model

The algorithm used in this thesis to calculate TEC from GPS measurements is described by Opperman (2007). The ASHA model is based on the Conventional Spherical Harmonic Analysis (CSHA) methodology described by Schaer (1999), which is originally a global implementation and was adapted for the South African region to be a regional model. The algorithm is designed to detect and remove or correct signal outliers and signal cycle slips in the preprocessing of the GPS data. In addition, the procedure also corrects for receiver and satellite (instrument) biases in the derived TEC data and only TEC observations with elevation angles above 20^o are used to avoid multipath effects. GPS basically provide measurements of slant TEC. However, for absolute TEC mapping using ground-based GPS data, the TEC along the vertical is of main interest. The slant TEC data are therefore mapped to the vertical by applying an elevation-dependent mapping function which is based on a single layer approximation at a 400 km height, given by:

$$STEC = \frac{VTEC}{\cos\chi'}, \;\; \sin\chi' = \frac{R_e}{R_e + r}\sin\chi \tag{5.2}$$

where χ is the satellite zenith angle, R_e is the Earth equatorial radius (6378.134 km), r is the assumed ionospheric shell height and χ' is the angle between the vertical and the signal path between the satellite and the receiver at the ionospheric pierce point. The algorithm fits the measurements to a spherical harmonic expansion in geographic latitude and sun-fixed longitude in order to solve the spatial and temporal variations of TEC in GPS data and to produce TEC maps over the entire region.

5.2.2 TEC variability

This section aims to highlight some general variabilities of TEC and references various studies that provided comprehensive investigations into the behaviour of TEC under different conditions. For example, Opperman (2007) and Habarulema *et al.* (2007) have undertaken extensive studies of TEC over South Africa. These and the many other groups such as Rama Rao *et al.* (2006), Yizengaw and Essex (2002) and Bagiya *et al.* (2009) provide information about the temporal and spatial variation of TEC.

5.2.2.1 Diurnal

The diurnal pattern of TEC exhibits a steady increase from about sunrise to an afternoon maximum and then falls to a minimum just before sunset. Figure 5.2 shows the diurnal pattern for a typical quiet day of 20th June 2005 over the Grahamstown ionospheric station. The diurnal characteristics of TEC have seasonal, solar activity, geomagnetic activity and latitudinal dependence.

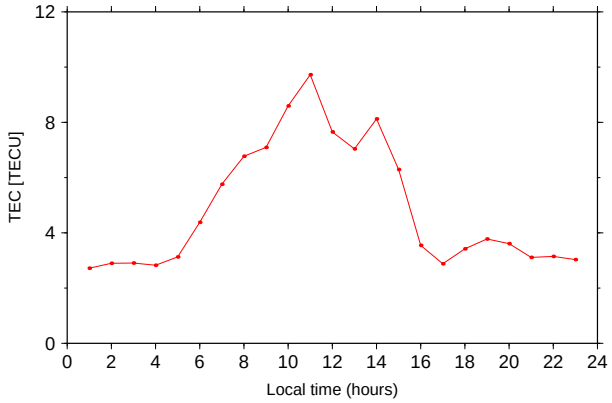


Fig. 5.2: Typical diurnal variation of GPS-TEC over the Grahamstown during winter

5.2.2.2 Seasonal

The noon values of TEC recorded at Grahamstown for the year 2005 are shown in Figure 5.3, which illustrates how TEC varies during different seasons.
The TEC values are high in equinox months, followed by more or less similar values in summer and winter.

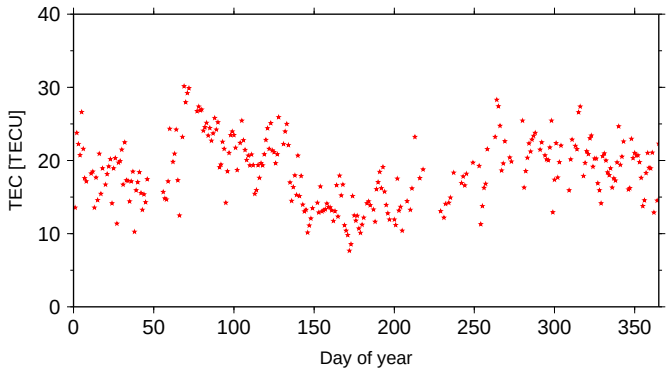


Fig. 5.3: Midday TEC values observed over Grahamstown in 2005. This indicates the typical seasonal TEC variation over Grahamstown during 2005

5.2.2.3 Geographical location

TEC data from various GPS stations over South Africa have been used to illustrate the horizontal (latitudinal as well as longitudinal) variation of TEC as shown in Figure 5.4.

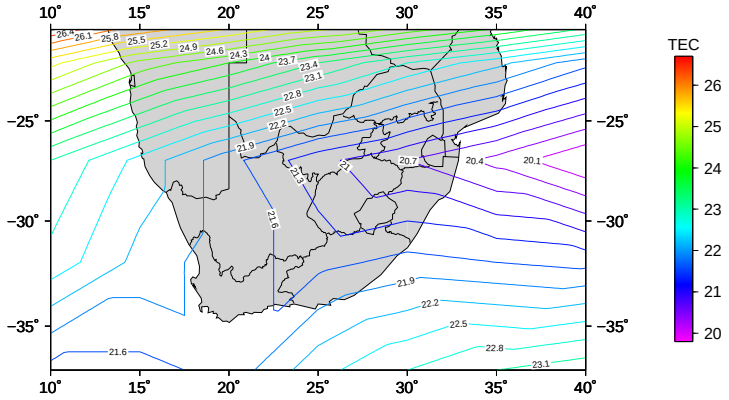


Fig. 5.4: Latitudinal and longitudinal variation of TEC over South Africa

5.2.2.4 Solar activity

Other studies investigated the effects of solar activity on TEC, for example, Warnant (2000) observed that TEC values become higher with increasing solar activity and Rama Rao *et al.* (2000) observed a direct correlation between solar activity and the ionisation level, with higher values during a high solar activity period and low value during a low solar activity period.

5.2.2.5 Geomagnetic activity

Many authors investigated the effects of geomagnetic storms on TEC. Jakowski *et al.* (1999) Muruyama *et al.* (2004), Munnucci *et al.* (2005) and others have reported both

positive and negative storm effects. Habarulema *et al.* (2009) provided a detailed study of storm time TEC variations over South Africa.

5.3 Vertical Ne profiling using GPS-TEC

The total electron content (TEC) of the ionosphere, being the line integral of the Ne encountered by a trans-ionospheric wave along its path of propagation, does not convey any information about the structures within the ionosphere.

Several studies devoted their efforts to develop techniques using GPS-TEC to provide vertical profiling of the ionospheric Ne. These techniques include:

(a) Ionospheric tomography, a technique for imaging the vertical cross section through ionospheric Ne. GPS-TEC values are used with tomographic reconstruction algorithms to reconstruct the ionospheric Ne for a relevant scenario (Sutton and Na, 1996).

(b) Ionospheric Radio Occultation (IRO), a situation which occurs when a transmitting GPS satellite, setting or rising behind the Earth's limb, is viewed by a Low Earth Orbiting (LEO) satellite from which vertical profiles of ionospheric parameters can be derived (Jakowski *et al.*, 2002).

(c) Stankov and Muhtarov (2001) proposed an approach that uses a combination of GPS-TEC measurements, ionosonde measurements and the upper transition height to reconstruct the Ne profile. The technique uses a profiler, such as Chapman, sech-squared or exponential, to construct a system of equations from which the unknown ion scale heights can be calculated. A unique Ne profile for the conditions specified by the GPS and ionosonde measurements can then be constructed.

5.3.1 Ionospheric radio occultation

It is not straight foward to establish altitude-dependent profiles of electron densities using ground-based GPS observations. However, such profiling can be provided by the dual-frequency phase measurements of a GPS receiver on a LEO satellite tracking a GPS satellite that is setting or rising through the Earth's ionosphere until it is occulted by the Earth's limb, as shown in Figure 5.5.

As the occultation occurs, the relative motion between the GPS and LEO satellites provides a vertical scanning of the ionosphere i.e. the radio waves sample successive layers of the ionosphere (Figure 5.6), providing vertical Ne profiles from the LEO orbit height down to the bottomside (Jakowski *et al.*, 2002).

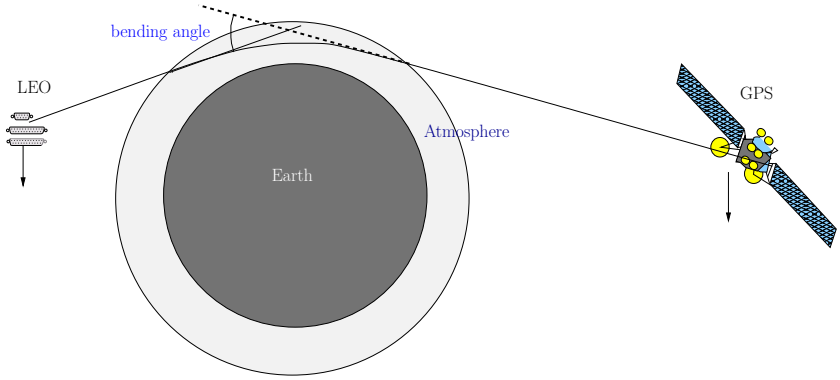


Fig. 5.5: Schematic view of a radio occultation event involving a GPS satellite and a LEO satellite

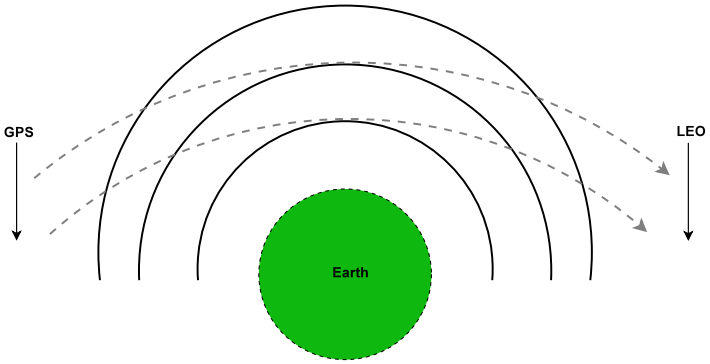


Fig. 5.6: Ionospheric radio occultation

During such occultation events, the GPS signal is both bent and slowed down, causing a delay in the arrival of the signal at the LEO. Combining these measurements with the satellites' position and velocity information, the small phase path increase due to the atmosphere during the occultation event can be derived. The phase variation of the GPS signal is then used to retrieve a vertical profile of ionospheric bending angles. Assuming a local spherical symmetric atmosphere, the vertical profiles of Ne can be derived. Ionospheric Radio Occultation (IRO) has become a powerful new technique to sense the vertical and horizontal structure of the ionosphere remotely with high data coverage on a global scale (Jakowski *et al.*, 2004). On a typical day with a full GPS constellation, occultation events occur frequently and a single LEO satellite in a polar orbit may observe several hundred globally distributed occultation events.

5.3.2 Ionospheric tomography

Ionospheric tomography involves reconstructing a 3-dimensional image of the Ne distribution from the TEC value, which is the integral of the image in a given direction (Yizengaw, 2004). Tomographic reconstruction techniques in general have broad applications, for example in the medical sciences (Defrise, 2001) where they are

used to produce high quality images of a section of the human body. They are also used in the mapping of underground resources (Faletič, 2005) such as the study of ocean structures and geological formations.

Austen *et al.* (1988) demonstrated that it was possible to use the tomographic reconstruction techniques to reconstruct the structure of the ionosphere successfully from TEC data. Many authors thereafter have further improved the technique and successfully applied it to various studies of the ionospheric structure both regionally and globally (see for instance Kunitake (1996), Ruzhin *et al.* (1998), Hansen (2002), Schlüter *et al.* (2002), Choi *et al.* (2006), Wen *et al.* (2007), Erturk *et al.* (2009))

Ionospheric tomography poses some physical limitations in the performance of the tomographic algorithms, due to a combination of factors including that the receivers are placed on a spherical Earth's surface at any possible location, not necessarily equidistant from each other. Furthermore, there is a limit to the observation angle and the number of projection samples that can be collected because the number of GPS satellites are limited and they trace a path over the receiver not necessarily designed for ionospheric tomography. Due to these limitations, for example, it is not possible to decompose its integral properly into the different values of Ne in the vertical. For this reason conventional tomographic imaging methods have to be modified to overcome these limitations. Tomographic methods, which include a priori information about the ionosphere, are developed where the vertical information is added to the method from a range of background ionospheres (for example, from a background model), representing many possible peak heights. In this way it is possible to decorrelate these layers and to obtain more realistic profiles. In these methods the ionospheric Ne is modelled as a linear combination of two dimensional basis functions. Generally, the discrete inverse theory approach is applied such that TEC along some path p_i represents a finite sum of shorter integrals along segments of the ray path length p_i, and is given by:

$$TEC = \sum_{j=1}^{M} n_j d_{ij} + e_j \tag{5.3}$$

where d_{ij} is a two dimensional basis function. The two dimensional basis functions are obtained as the product of the vertical and horizontal basis functions. Either Legendre or Fourier polynomials are used as horizontal basis functions (Blanch, 2003; Sutton and Na, 1994) and the background ionospheric models provide the vertical basis functions. The background ionospheric model could be, for example, an ionosonde measurement or a Chapman layer profile.

5.4 Combining ionosonde and GPS data

As mentioned in previous sections, ground-based GPS-TEC measurements are basically insensitive to the vertical distribution of the Ne since the geometry of the ground-based GPS observations is mainly vertical, and therefore it is difficult to distinguish the contribution of each layer to the vertical distribution.

Therefore, an additional complementary data source is needed to decorrelate the layers in the vertical direction. In this context, Stankov and Muhtarov (2001) proposed an approach to use a combination of GPS-TEC data and ionosonde measurements to perform the vertical reconstruction of the Ne profile. Continuing the idea of combining different kinds of data, this chapter is focused on mixing ground based GPS-TEC measurements, ionosonde data and modeled UTH values with the objective of providing an alternative solution to identify the vertical distribution of the Ne. The method is implemented for a single station, Grahamstown (33.3°S, 26.5°E), South Africa, where a GPS receiver and an ionosonde are co-located. Preliminary results for the implementation which used the Epstein function (equation 2.2) to approximate the individual ion density distributions in the topside ionosphere, are presented. A key assumption is made that the ionosphere is statistically neutral and that the ions are singly charged, therefore the Ne is given as the sum of the individual ion densities. Instead of approximating the Ne directly, the densities of the constituent ions which have a major influence on the distribution of the Ne in the topside ionosphere, are approximated individually, from which the Ne distribution is derived. This allows for the inclusion of other important profile shape factors such as the UTH, which simplifies the problem.

5.4.1 Input data for use in the reconstruction procedure

5.4.1.1 GPS-TEC measurements

The GPS-TEC values were derived using the Adjusted Spherical Harmonic Analysis (ASHA) model described in Opperman (2007), and reviewed briefly in section 5.2.1. The ASHA model is an implementation of the Conventional Spherical Harmonic Analysis (CSHA) methodology (Schaer, 1999), adjusted in order to adapt it for a regional application to take advantage of densely distributed observations in a relatively small geographic region.

5.4.1.2 Ionosonde measurements

The ionosonde measurements used were from the Grahamstown (33.3°S, 26.5°E) ionospheric station, which is co-located with the GPS receiver, an important requirement for this approach. The required ionosonde parameters are the peak Ne ($NmF2$), the height of the peak Ne (h_mF2) and the bottomside ionospheric TEC (TEC_b). TEC_b was determined from the bottomside Ne profile by integrating the ionosonde Ne measurements up to h_mF2. The vertical GPS-TEC measurements can then be split in two contributions, one due to the bottomside ionosphere (TEC_b) and the other due to the topside ionosphere (TEC_t), so that TEC_t is given by:

$$TEC_t = GPS\text{-}TEC - TEC_b \tag{5.4}$$

5.4.1.3 Upper transition height

Another key to this approach is the knowledge of the upper transition height (UTH), to provide information about the relative composition of the ions, which is an important factor affecting the distribution of Ne in the topside ionosphere. The UTH is always above the F2-peak, and therefore this height is useful as a reference point to anchor the Ne profile. The values of the UTH were determined from two sources: (1) The Field Line Interhemispheric Plasma (FLIP) model (Richards *et al.*, 2000), a theoretical model that solves the continuity and momentum equations in one dimension to calculate the plasma densities for the O^+, H^+, He^+ and N^+ ions, and the energy equations for electron and ion temperatures along a magnetic flux tube from 80 km in the northern hemisphere through the plasmasphere to 80 km in the southern hemisphere (Richards, 2001). Apart from the ion densities, the model also provides the transition heights for the O^+ - H^+ ions. (2) The second source was the neural network model developed in this study and described in chapter three. The neural network model is based on available topside sounder data. The UTH values from this (empirical) model provide an additional advantage in that the reconstructed Ne profiles are tied to additional measurements upon which the model is based.

5.4.2 Determining the profile function

Assuming that helium ions have little effect on the Ne profiles under most conditions (Carlson and Gordon, 1966), their presence can be neglected. Therefore, the major ion species present in the topside ionosphere are the hydrogen and oxygen ions (Titheridge, 1972). The topside ionospheric Ne can be expressed as a sum of the constituent O^+ and H^+ ion density profiles:

$$N_e(h) = N_{O^+}(h) + N_{H^+}(h) \tag{5.5}$$

The Epstein functions are used to analytically approximate the individual ion density distributions (N_{O^+} and N_{H^+}), thus the reconstruction formula for the Ne as a function of altitude (h) is given by:

$$N_e(h) = N_{O^+}(h_mF2)sech^2\left(\frac{h - h_mF2}{2H_{O^+}}\right) + N_{H^+}(h_mF2)sech^2\left(\frac{h - h_mF2}{2H_{H^+}}\right) \tag{5.6}$$

where $N_{O^+}(h_mF2)$ and $N_{H^+}(h_mF2)$ are the respective oxygen and hydrogen densities at the F2-peak, H_{O^+} is the oxygen scale height and H_{H^+} is the hydrogen scale height. Since h_mF2 is obtained from the ionosonde measurements, equation 5.6 has four unknowns namely: $N_{O^+}(h_mF2)$, $N_{H^+}(h_mF2)$, H_{O^+} and H_{H^+}.

Based on the theoretical definition of the plasma scale height, the H_{O^+} can be expressed in terms of H_{H^+}, reducing the number of unknowns to three. Following the theoretical definition of the plasma scale height, (equation 2.5), under diffusive equilibrium along magnetic field lines and isothermal conditions in the topside ionosphere the ratio H_{H^+}/H_{O^+} of the hydrogen and oxygen scale heights is approximately equal to their mass ratio ≈16 (Kutiev and Marinov, 2007). Thus, the H_{H^+} can be expressed as:

$$H_{H^+} = 16H_{O^+} \tag{5.7}$$

It is important to emphasise that this holds only if scale heights are calculated along magnetic field lines. However, in this case the scale heights are in the vertical direction, thus the ratio does not stay the same (Kutiev and Marinov, 2007). A correction factor based on simple geometrical considerations of the geomagnetic field was used to map the scale heights along magnetic field lines on to the vertical direction, which will in effect distribute the plasma density in the vertical direction (Stankov *et al.*, 2003). The derivation for the correction factor is described in section 4.5, and is given as:

$$\tau = \sin(\arctan[2\tan(\lambda)]) \tag{5.8}$$

Applying 5.8 in 5.7 and substituting into equation 5.6 yields

$$N_e(h) = N_{O^+}(h_mF2)sech^2\left(\frac{h-h_mF2}{2H_{O^+}}\right) + N_{H^+}(h_mF2)sech^2\left(\frac{h-h_mF2}{32\tau H_{O^+}}\right) \quad (5.9)$$

Integrating equation 5.9 from h_mF2 to infinity yields equation 5.10 (Stankov and Muhtarov, 2001; Stankov *et al.*, 2002) for TEC_t, the contribution of the topside ionosphere to TEC:

$$TEC_t = 2H_{O^+}N_{O^+}(h_mF2) + 32H_{O^+}N_{H^+}(h_mF2) \quad (5.10)$$

The upper transition height and the F2-peak provide reference points to anchor the profile and simplify the reconstruction problem.

- At the upper transition height (UTH) the O^+ and H^+ ion densities are equal, a condition represented by:

$$N_{O^+}(h_mF2)sech^2(\frac{UTH-h_mF2}{2H_{O^+}}) = N_{H^+}(h_mF2)sech^2(\frac{UTH-h_mF2}{32\tau H_{O^+}}) \quad (5.11)$$

  The UTH lies in the transition region from a predominantly O^+ topside ionosphere to a predominantly H^+ plasmasphere. The Ne profile changes its slope in this region and the different scale heights of the constituent O^+ and H^+ ion profiles causes the gradient of the vertical Ne distribution to increase sharply. This height can serve as a base for finding the relative quantity of H^+ and O^+ ions in the topside ionosphere.

- At the F2 layer peak, the sum of the O^+ and H^+ ion densities is equal to the measured peak Ne ($NmF2$) following the condition of quasi-neutrality. This condition is shown in equation 5.12:

$$NmF2 = N_{O^+}(h_mF2) + N_{H^+}(h_mF2) \quad (5.12)$$

Equations 5.10, 5.11 and 5.12 form a system of three equations with three unknown parameters namely: $N_{O^+}(h_mF2)$, $N_{H^+}(h_mF2)$ and H_{O^+}. Three key inputs ($NmF2$, TEC_t and the UTH) are determined from the data sources described in section 5.4.1. This system of equations was solved numerically using a Newton iterative procedure described by Yang *et al.* (2005) to obtain the O^+ ion scale height. The calculated scale height is then used in the reconstruction equation 5.9 to obtain the Ne as a function

of height. The topside Ne profiles constructed using this approach are presented in the next section.

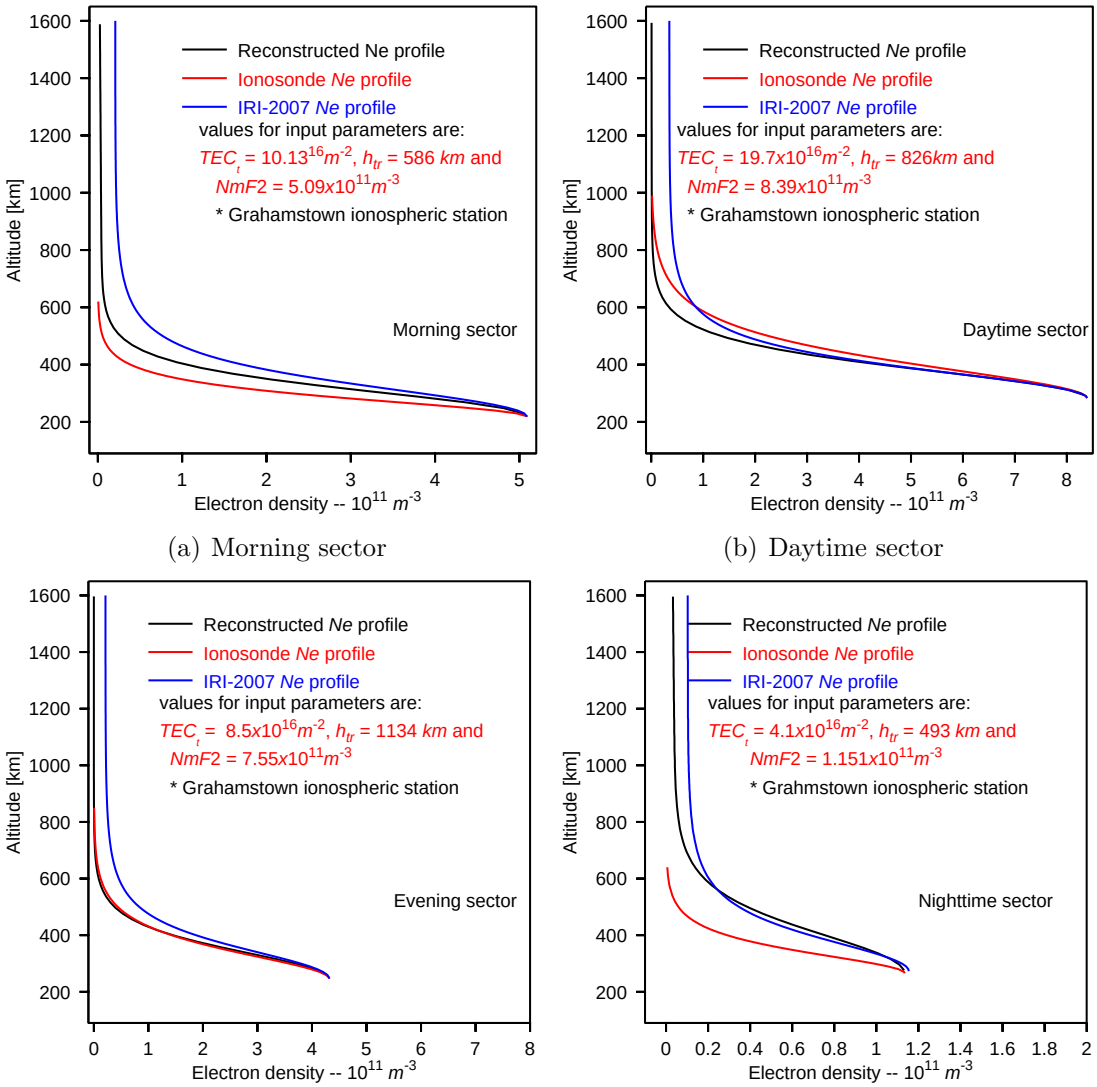
5.5 Results and analysis

As noted earlier, an essential requirement for the procedure is an ionosonde and a GPS receiver that are co-located. The procedure was implemented and this section presents the results of the implementation. It was desirable to test the procedure on actual measurements, however there were no measured data to compare with the reconstructed profiles. Instead a comparison with other models was undertaken in order to get a general overview of how the procedure performs against the commonly used methods to predict the topside ionosphere over the South African region. Two models, the IRI-2007 model (Bilitza, 1990) and the topside model used in the ionosonde scaling software (Huang and Reinisch, 1996), were used for comparison with the reconstructed profiles. This comparison does not give an indication of the accuracy of the approach since it is a comparison of a model with other models. The only benefit of this method of comparison, and the reason it is presented here is to indicate that the proposed approach produces reasonable profiles as determined by the global models already in operation.

The described reconstruction procedure was performed using the data from the Grahamstown (33.3°S, 26.5°E) ionospheric station producing the Ne as function of altitude for different scenarios. The GPS receiver at the Grahamstown station was installed at the beginning of 2005, therefore the GPS data is only available from then onwards. The reconstruction was done for different local time sectors: morning (06h00 LT), daytime (12h00 LT), evening (18h00 LT) and nighttime (00h00 LT) on 4 April 2005. The resulting Ne profiles (black curves) compared with the ionosonde profiles (red curves) and the IRI-2007 profiles (blue curves) are presented in Figure 5.7. The UTH values were provided by the FLIP model (Richards, 2001).

Similarly, Figure 5.8 shows the reconstruction for the same scenarios presented in Figure 5.7 representing morning, daytime, evening and nighttime during the course of the same day, but with the UTH values provided by the empirical neural network based model developed in this study, and described in chapter three.

The topside profiles provided by the ionosonde are not actual measurements, but an approximation derived with the α-Chapman function using a scale height deduced from the bottomside profile (Huang and Reinisch, 1996; Reinisch and Huang, 2001). The IRI model, on the other hand, is based on the Alouette and ISIS topside sounder data



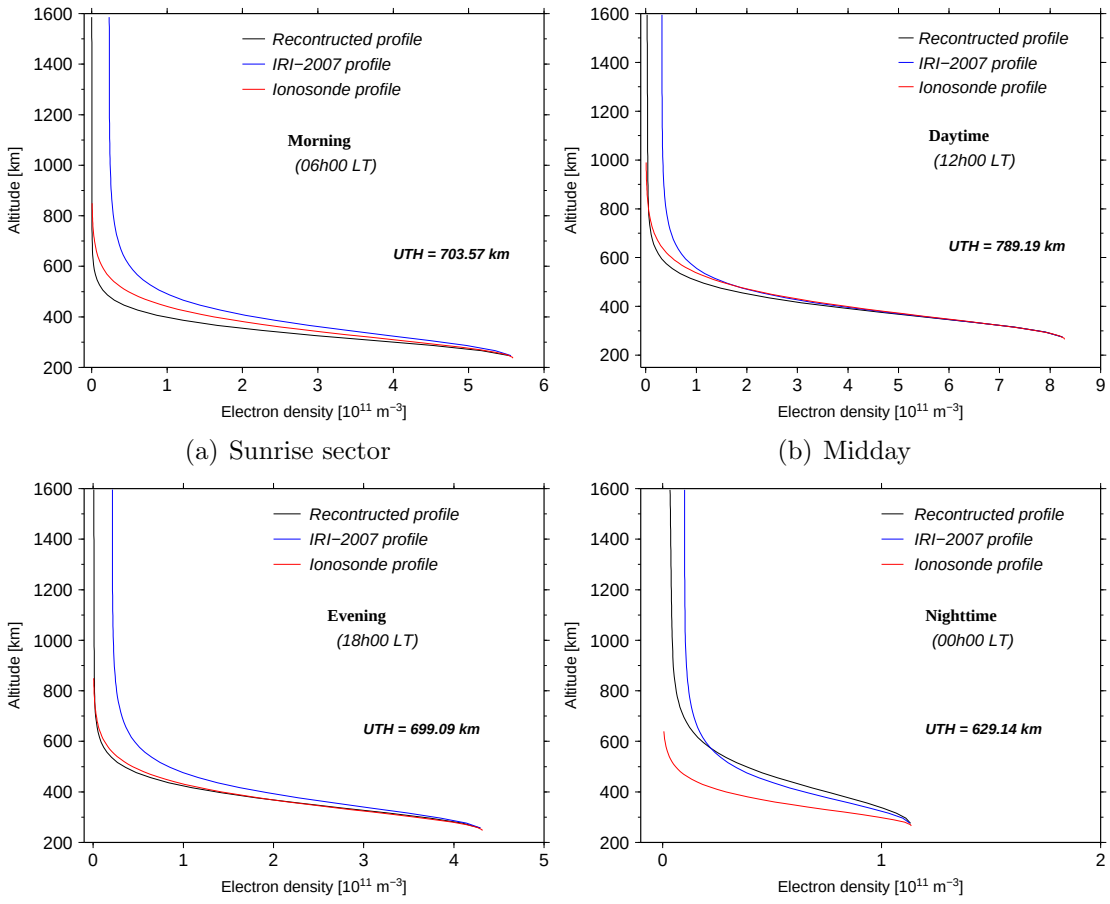
(a) Morning sector (b) Daytime sector

(c) Evening sector (d) Nighttime sector

Fig. 5.7: Reconstructed topside Ne profiles for morning (a), daytime (b), evening (c) and nighttime (d) sectors compared with the corresponding ionosonde profiles and the IRI model profiles. Values for the input parameters used are indicated in each plot. The UTH values where calculated with the FLIP model

as explained in section 2.2.2.2. Both the ionosonde and the IRI profiles are known to have shortcomings, the comparison therefore only provides a general picture of how the procedure performs and does not indicate the accuracy of the method. To quantify the accuracy of the method would require actual measurements of the topside ionosphere.

Figure 5.9 shows the results for the Ne reconstruction performed for midday local time on different days of the year (82, 171, 263 and 343), representing the different seasons autumn, winter, spring and summer respectively.



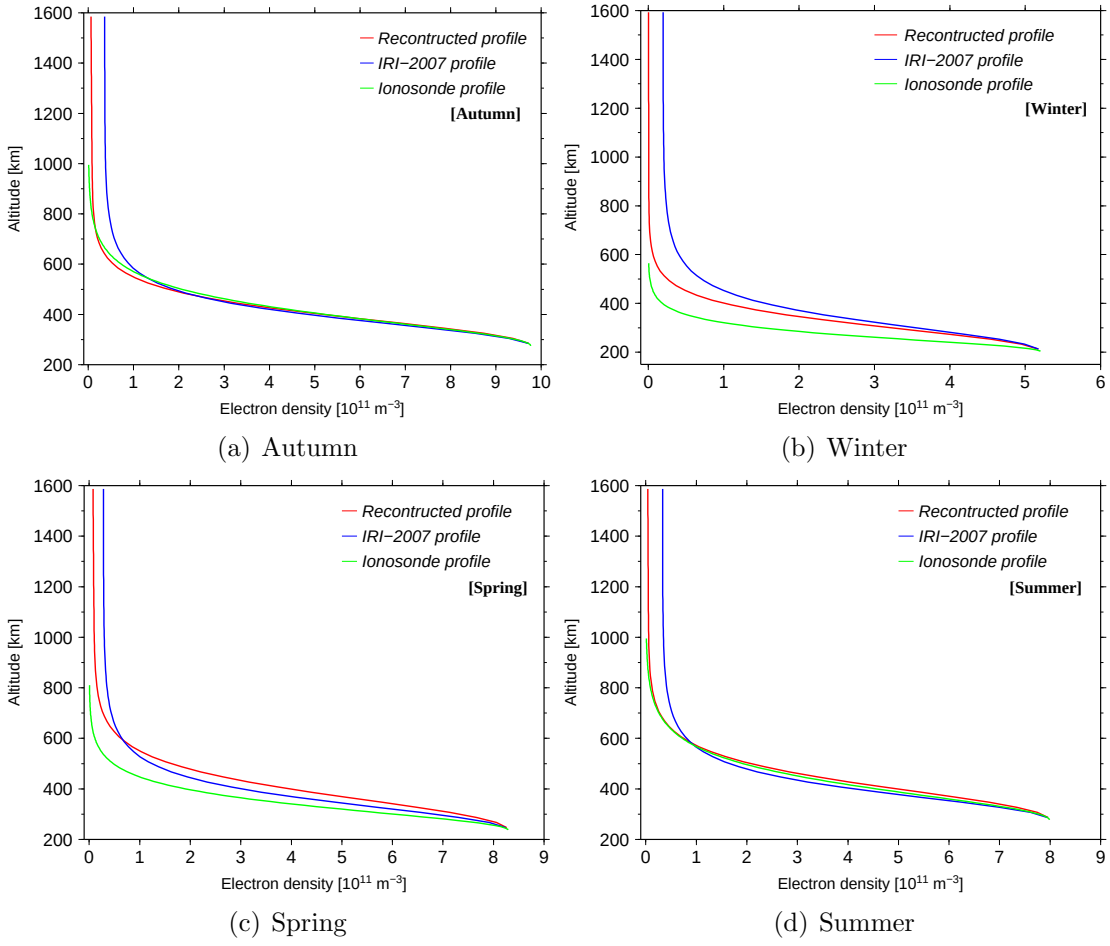
(a) Sunrise sector (b) Midday

(c) Sunset sector (d) Midnight

Fig. 5.8: Reconstructed topside Ne profiles for morning (a), daytime (b), evening (c) and nighttime (d) sectors compared with the corresponding ionosonde profiles (red curves) and the IRI model profiles (blue curves). The empirically obtained UTH values are indicated on each plot

The focus here is on the vertical structure of the Ne. The diurnal, seasonal, latitudinal, solar activity and geomagnetic activity response of the reconstructed profiles is contained within the inputs (GPS and ionosonde measurements and the UTH values) which are specific for a given set of geophysical conditions, and thus providing a unique Ne profile for each scenario.

These results show that the Ne profile can be reconstructed from its integral quantity, TEC, showing the smooth and continuous decrease of the Ne with altitude comparable with other empirically obtained profiles. This approach offers an opportunity to improve topside modelling efforts and provide valuable information about the topside ionosphere, a region that is difficult to model due to the scarcity of measured data. The approach has the advantage that the constructed profile is tied to reliable measured TEC values offering a higher level of confidence in the resulting Ne.



(a) Autumn (b) Winter

(c) Spring (d) Summer

Fig. 5.9: Midday (LT) reconstructed topside Ne profiles (red) for autumn (a), winter (b), spring (c) and summer (d) compared with the corresponding ionosonde profiles(red) and the IRI model profiles (blue) for the dates sampled.

5.6 Summary and discussion

The work in this chapter demonstrates that while TEC simply represents the integral of the Ne along the path, it can be useful in providing the information about the spatial variation of Ne along the path caused by the irregular structures in the ionosphere. Based on the technique by Stankov and Muhtarov (2001), the combination of the GPS and ionosonde measurements at Grahamstown (33.3°S, 26.5°E), South Africa was used to reconstruct the vertical Ne profile over this location. The results presented show that this approach can be a useful tool in characterising the topside ionospheric Ne over the South African region, where measured data is sparse and does not properly represent the various geophysical conditions. Using the data from the four ionosondes in South Africa and the dense network of GPS receivers shown in Figure 5.1, the procedure can be expanded in longitude and latitude to provide a mapping of the entire region.

As mentioned before, the Epstein function was used for this study. From Figure 2.2, it can be seen that the α-Chapman function would produce the largest TEC values since it gives the highest Ne values at all altitudes, and the exponential function is the lowest for the same scale height. The Epstein function provides a fair average between the extremes. For a mid-latitude region such as the South African region, where the ionospheric parameters behave in a rather predictable manner, the Epstein function can be appropriate. However, until proper tests are undertaken to establish which function provides the best performance and is most appropriate for application in this region, these results should be viewed as a first attempt at implementing the approach. It should be emphasised again that the success of the approach and determination of which function provides the best performance can best be determined when simultaneous topside Ne and GPS measurements together with the ionosonde measurements are available.

Single values of the ion scale heights are determined across the entire altitude range of the topside ionosphere and plasmasphere. Thus the ion scale heights are assumed to be constant throughout the topside ionosphere and plasmasphere. However, as demonstrated in Figure 4.3, the ion scale heights vary with altitude. This follows from the theoretical definition of the plasma scale height (equation 2.5) which depends on the ion and electron temperatures and the masses of the constituent ions. The electron and ion temperatures in the topside ionosphere increases with altitude (for example Figure 4.1), thus the plasma scale height also increases (Kutiev and Marinov, 2007). This approach, therefore, does not take into account the altitude variation of the scale height. The scale height plays a key role in determining the shape of ionospheric Ne profiles using the 'profiler' functions (equations 2.1, 2.2, 2.3, 2.4), therefore, if the scale height is incorrectly determined, this would affect the shape of the calculated profile. A more accurate approximation of the topside profile would require the construction of a suitable scale height function that represents the altitude variation of the scale height.

Future work should involve applying this technique at a location with an ionosonde co-located with a GPS receiver, including measured topside profiles to validate this topside reconstruction technique. This would show how the calculated profiles compare with actual measurements. Should it become possible to validate this technique with topside measurements, the procedure could be refined for use over African regions which currently experience a lack of ionospheric data, but are likely or are planning to install GPS receivers.

GPS based techniques offer a promising tool for ionospheric characterisation. They

provide a combination of dense global coverage, and the capability to give continuous measurements of TEC values, as well as being a relatively inexpensive technique.

Chapter 6

Discussion and Conclusions

6.1 General remarks

Modern society have become increasingly reliant on technological systems such as satellite operations, telecommunication and navigation systems whose performance and reliability depends on the state of the ionosphere. Therefore, proper understanding of ionospheric dynamics has become of fundamental practical importance. In particular, the Ne is an ionospheric variable of great importance for systems that use trans-ionospheric radio waves because most of the ionospheric effects on the radio signals are directly proportional to the Ne. If modeled properly, the ionospheric Ne can be useful for obtaining corrections for affected systems. The goal of scientific research in this area is to provide a deeper understanding of the characteristics and consequences of ionospheric structures and dynamics (Cander, 2008). A wide range of models have been developed in this regard, to assist the various ionospheric applications. The topside ionosphere is an important component of the whole ionospheric dynamic system, and due to the difficulty with measuring this region, it is not as well understood as the bottomside ionosphere. This thesis provided a detailed overview of the various international efforts to model the topside ionosphere highlighting the broad international participation, with a specific emphasis on the implication of these efforts to the understanding of the topside over the South African region.

The goal of this thesis is to provide a comprehensive understanding of topside ionospheric modelling over South Africa, which will serve as a starting point towards the development of a topside ionosphere model that is suitable for this region. Such a model would eventually be fitted to the South African Bottomside Ionospheric model (SABIM) which is currently valid only for the bottomside ionosphere. A number of steps were taken to address the problem and achieve the intended results. In the process, several challenges were identified and need to be addressed.

Steps taken

1) First, there was a requirement for a holistic understanding of the available topside ionosphere measurements in terms of how much data is available and how the data is distributed over the various geophysical conditions, with a specific emphasis on the coverage of the South African region. This investigation revealed that the available topside ionospheric Ne measurements are limited and the data irregularly sampled over the relevant conditions. In particular, the South African region is not well covered within the data sets.

2) The second requirement was for a proper understanding of the various modelling techniques available in order to identify the approach that is most suitable for modelling the South African topside ionosphere. A wide range of modelling approaches have been implemented by various authors in efforts to describe the global topside ionospheric behaviour. This thesis presented a detailed overview of the various topside modelling techniques providing the information required for the topside modelling efforts in South Africa.

3) Three different techniques, one based on the empirical approach, the other based on the theoretical approach and the third based on data assimilation techniques involving GPS observations were identified and reviewed for possible application in the South African efforts to model the topside ionosphere.

In the sections that follow, a summary is provided of the different contributions of this thesis to achieve the goal and to address the challenges identified.

6.2 Summary of results

6.2.1 Empirical approach

The first step (Chapter 3) addressed the availability of measured data and reviewed progress towards the empirical characterisation of the vertical Ne distribution in the topside ionosphere. Many of the empirical topside ionospheric models reported in the literature were constructed using functional equations. Such techniques are often limited by approximations and assumptions of linearity, normality and variable independence, because the underlying structural relations cannot be easily represented using mathematical formulae in most complex dynamic systems such as the highly non-linear ionosphere (Thidé, 2007). In this thesis, Artificial Neural Network techniques (ANNs) were employed. ANNs are designed to capture many kinds of relationships

between the variables and can provide an analytical alternative to conventional techniques.

The ionospheric parameters, H_{O^+}, H_{H^+} and UTH , were modeled as functions of the the day of year, local time, geographic latitude and longitude, magnetic inclination, 12-month running mean of the sunspot number, height of the F2-peak (h_mF2) and peak Ne ($NmF2$), from which the Ne distribution with height can be derived. The model variables (H_{O^+}, H_{H^+} and UTH) did not show identical features to those reported in other studies. However, the resulting Ne profiles constructed with a combination of the variables were comparable with measured profiles, and also with the profiles predicted by the IRI model. But comparison of the model Ne results with the measured profiles and with the IRI profiles showed that the ANN approach did not provide a significant improvement compared to the IRI model results.

Limitations

- ANNs rely on the amount of data they are presented with from which they can learn the pattern of variations. If presented with insufficient data they cannot capture the complex ionospheric variabilities correctly. For example, for the UTH it was observed that, in general, the neural network results did not capture the large peaks seen in the measurements. Instead the ANN output is averaged to a narrow range of between 500 km and 1100 km. It was, therefore, concluded that although ANNs are capable of modelling phenomena which otherwise may have been very difficult or impossible to represent with functional equations, sometimes they deliver only minimal performance. Without a more suitable representative database it is not possible to use this technique optimally.

- The major disadvantage with empirical models is that they generally represent average conditions and are useful in predicting monthly median values of basic ionospheric parameters for a given set of geophysical conditions (Sibanda and McKinnell, 2009b). There is, therefore, a limit to the accuracy with which they may represent the instantaneous ionosphere. In addition, they rely on good quality measured data, which for the case of the topside ionosphere, is not easy to collect. However, even with all these limitations, empirical models are probably the most practical for ionospheric applications, especially in those areas sufficiently covered by observations since they are based on measurements.

- The general paucity of measured topside data leads to the problem that the empirical approach based on these data is limited in representing the actual

conditions in the ionosphere, especially in the regions that are not well covered.

6.2.2 Theoretical approach

Chapter 4 focused on the Ne estimation method based on theoretical considerations of the physical processes that play a significant role in controlling the distribution of charged particles in the topside ionosphere. An ionospheric estimation algorithm was designed based on the well known diffusive equilibrium theory. The application of the diffusive equilibrium approach requires the knowledge of various ionospheric parameters which include: electron and ion temperatures (T_e and T_i), the number densities (N_{H^+} and N_{O^+}) of the constituent ions and the number densities of the neutral particles (N_H and N_O) at the starting height. The preliminary results demonstrated the capabilities of the diffusive equilibrium approach to provide an option for modelling the topside ionopshere, especially in the regions where there are not enough measurements. Theoretical models have the potential to be another option for topside modelling and can provide an understanding of the natural variabilities.

Limitations

- The greatest disadvantage with this approach is its dependence on several ionospheric parameters as inputs, which may not be well known. For this reason, it was concluded that, this option is currently inadequate for the South African ionosphere since the required input parameters (electron and ion temperatures, neutral temperatures and relative abundances of the constituent ions) are not well studied in the region. This approach would be a more viable option if measurements such as GPS-TEC, which are more readily available could be used instead. The incorporation of such data should be explored and investigated further.

- The theoretical approach also has limitations which are mainly due to the usual assumptions and approximations.

6.2.3 GPS based approach

There has been a growing interest in developing models involving data assimilation techniques in order to use the much available real-time GPS measurements, a source of data that is continually expanding. In chapter 5 it was demonstrated that the vertical distribution of the ionospheric Ne can be reconstructed from its integral, the TEC. The chapter explored the capabilities of using the GPS-TEC measurements with other ionospheric measurements to reconstruct the vertical structure of the topside ionospheric

Ne. A method that employs a combination of GPS-TEC, ionosonde measurements and modeled UTH (proposed by Stankov and Muhtarov (2001)) to reconstruct the topside Ne profile was successfully implemented. This relatively new approach presents the potential for improving the studies of the topside ionosphere over the South African region significantly. The advantage with this approach is that the constructed profile represents instantaneous ionospheric conditions since it is constructed from the observed ionosonde and GPS observations. Simultaneous GPS and ionosonde measurements at Grahamstown became available only in 2005 when a GPS receiver was installed at the Grahamstown ionosonde station. However, there were no measured topside profiles to compare with the results. It was therefore difficult to determine the accuracy of the method in terms of how it compares with actual measurements.

Limitations

- One challenge with this approach is the unavailability of UTH measurements. Instead the UTH values are obtained from models. Two models were tested in this thesis, (i) the FLIP model, a theoretical model, and (ii) the empirical neural network based model discussed in chapter three. For the examples shown, it was observed that using the empirically obtained UTH values did not provide any significant improvement compared to using the theoretically obtained UTH values.

- Another limitation of this approach is that single values of the ion scale heights are obtained for each profile, which means the scale heights are assumed to be constant in the entire altitude range of the topside ionosphere and plasmasphere. However, this assumption is not entirely correct since the scale heights are functions of the ion and electron temperatures which vary with altitude (Kutiev and Marinov, 2007). The scale height is a key parameter in determining the shape of the profile, therefore, if it is not correctly obtained, the accuracy of the resulting profile would be compromised.

6.3 Recommendations for future work

This thesis encountered many problems which are not answered as yet, and more research work is required to address these challenges. In addition, the results presented in this thesis open several opportunities for future research projects. Some recommendations for future research directions are considered below.

1. In the case of the empirical approach presented in chapter three, only data from the Alouette and ISIS satellites were used. The Japanese ISS-b, the Russian Intercosmos-19 and Cosmos-1809 data sets were not included in this analysis.

- Future work should include the entire range of data sets listed in table 2.1 in order to expand the coverage of the data.
- Another area of future research on this approach is to explore ways of updating the developed neural network model with other data sources, such as the real-time GPS measurements, to adapt it to specific ionospheric conditions. It is expected that the success of such a process would result in a significant improvement.

2. With the imminent establishment of a South African space agency that will coordinate and facilitate the development of space missions in South Africa, the author recommends the inclusion of a topside ionospheric sounder on-board a satellite that will be within orbits that include South Africa. This will provide the long desired topside ionosphere data set for the South African region including parts of Africa, thereby enabling more research that will bridge the existing gap in terms of understanding, modelling and explaining the ionospheric variations over South Africa. The main challenge is the expense associated with topside sounder missions. However, the advantages, apart from providing the much needed absolute measurements for the South African region, includes enhancing and broadening the satellite training program in terms of skills and capacity building since it will open many opportunities for promissing scientific research projects.

3. The approach involving the use of GPS and ionosonde measurements was not tested against actual measurements. Efforts to find a location that has measured topside profiles as well as GPS and ionosonde measurements were not successful at the time of preparation of this thesis. However, until this approach is tested on actual measurements, it is not possible to determine the accuracy and reliability of the technique.

- The focus of the future work should continue to search for such simultaneous measurements in order to validate this complex reconstruction technique and demonstrate its capabilities and suitability for operational applications.
- Further, the use of constant scale heights in the topside ionosphere and plasmasphere means that the approach is not accounting correctly for the actual conditions in the topside ionosphere. An ideal estimation algorithm would involve scale heights that vary with altitude. Future work, therefore should explore ways to explicitly account for the altitude variation of the scale heights, instead of relying on the approximations as used in this work.

4. Finally, due to the limitations of each of the approaches reviewed, as pointed out in section 6.2, it was observed that at this stage characterisation of the topside ionosphere over the South African region cannot be properly achieved by a single approach among those discussed in this thesis. In order to achieve more satisfactory results, this thesis recommends that a combination of all or some of these techniques is required. Thus, an approach incorporating both experimental evidence from more readily available measurements such as GPS-TEC and theoretical considerations of the complex interplay between the ionospheric constituents be explored further. Success of such efforts will be important not only within the framework of ionospheric vertical Ne characterisation, but would also provide a deeper understanding of the response of the system to magnetospheric phenomena such as ionospheric and geomagnetic storms. It is therefore, foreseen that in the short term an approach to topside ionospheric modelling over South Africa should include both an empirical and a theoretical component, until such time as more measured data becomes available.

6.4 Concluding remarks

This thesis has provided a deeper understanding of the challenges of modelling the topside ionosphere over South Africa. Details are given of the efforts made to identify approaches and algorithms to follow in creating a model suitable for the South African region. Three different approaches were implemented but did not produce significant improvements compared to the results from the IRI model. It must be emphasised that no single approach of characterising the topside ionosphere among those discussed in this thesis is universal, each type has advantages as well as limitations, as discussed. However, the reconstruction technique that uses GPS-TEC with ionosonde measurements, discussed in chapter five, has the potential to be useful for the reconstruction of the topside ionosphere over the South African region. The combination of the wide coverage of GPS observations in South Africa and the versatility of the technique opens up the possibility for mapping the topside ionosphere over the entire South African region and gaining further understanding of the topside ionospheric dynamics in this region. The point is made that ultimately future long term planning should include a topside sounder mission for South Africa that will provide the measurements required to acquire a deeper understanding of the dynamics of this region.

References

Ajith, A., 'Articial Neural Networks', *in Handbook of Measuring System Design*, edited by Sydenham, P. H. and Thorn, R., pp. 901–908, John Wiley & Sons, Ltd., Oklahoma State University, Stillwater, Oklahoma, USA, 2005.

Anderson, D. N., 'Global ionospheric modelling', *in Modern Radio Science 1993*, edited by Matsumoto, H., p. 159, Oxford University Press, 1993.

Austen, J. R., Franke, S. J. and Liu, C. H., 'Ionospheric image using computerised tomography', *Radio Science*, 23: pp. 299–307, 1988.

Bagiya, M. S., Joshi, H. P., Iyer, K. N., Aggarwal, M., Ravindran, S. and Pathan, B. M., 'TEC variations during low solar activity period (2005–2007) near the equatorial ionospheric anomaly crest region in India', *Annales Geophysicae*, 27: pp. 1047–1057, 2009.

Banks, P. M., Schunk, R. W. and Raitt, W. J., 'The topside ionosphere - a region of dynamic transition', *Annual Review of Earth and Planetary Sciences*, 4: pp. 381–440, 1976.

Bassiri, S. and Hajj, G. A., 'Modeling the Global Positioning System signal propagation through the ionosphere', Telecommunications and Data Acquisition (TDA) progress report 42-110, 92-103, 1992.

Bauer, S. J., 'On the electron density distribution above the F2 peak (revised)', Technical Report D-1171, NASA, Washington, USA, 1962a.

Bauer, S. J., 'On the structure of the topside ionosphere', *Journal of the Atmospheric Sciences*, 16: pp. 276–278, 1962b.

Bauer, S. J., 'Diffusive equilibrium in the topside ionosphere', *in proceedings of the IEEE*, volume 57, pp. 1114–1118, 1969.

Benkova, N. P., 'Longitudinal features shown by topside sounder data and their importance in ionospheric mapping', *Advances in Space Research*, 10(8): pp. 57–66, 1990.

Benson, R. F., Reinisch, B. W., Green, J. L., Fung, S. F., Calvert, W., Haines, D. M., Bougeret, J. L., Manning, R., Carpenter, D. L., Gallagher, D. L., Reiff, P. and Taylor, W. W. L., 'Magnetospheric radio sounding on the IMAGE mission', *Radio Science Bulletin*, 285: pp. 9–20, 1998.

Bent, R. B., Llewellyn, S. K. and Schmid, P. E., 'Description and evaluation of the Bent ionospheric model', *in Space and Missile Syst. Org*, volume 1-3 of *Rep. SAMSO TR-72-81*, Natl. Tech. Inf. Serv., Springfield Va., 1972.

Bilitza, D., 'Implementation of the new electron temperature model in IRI', *Advances in Space Research*, 5(10): pp. 117–121, 1985.

Bilitza, D., 'International Reference Ionosphere', Technical report, NSSDC 90-92. WDC-A R&S, Greenbelt, Maryland, U.S.A, 1990.

Bilitza, D., 'Electron and ion temperature data for ionospheric modelling', *Advances in Space Research*, 11(10): pp. 139–148, 1991a.

Bilitza, D., 'Solar-terestrial models and application software', *Planetary Space Science*, 40(4): pp. 541–579, 1991b.

Bilitza, D., 'Topside models: status and future improvements', *Advances in Space Research*, 14: pp. 17–26, 1994.

Bilitza, D., 'Density profiles from the Aloutte 1, 2 and ISIS 1, 2 topside sounder instrument on CD-ROM and www', International Reference Ionosphere Newsletter, Vol 8, No. 1/2, 2001.

Bilitza, D., 'A correction for the IRI topside electron density model based on Alouette/ISIS topside sounder data', *Advances in Space Research*, 33: pp. 838–843, 2004.

Bilitza, D., Huang, X., Reinisch, B. W., Benson, R. F., Hills, H. K. and Schar, W. B., 'Topside ionogram Scaler with True Height Algorithm (TOPIST): automated processing of ISIS topside ionograms', *Radio Science*, 39(RS1S27): pp. 1–7, 2004.

Bilitza, D., Reinisch, B. W., Radicella, S. M., Pulinets, S., Gulyaeva, T. and Triskova, L., 'Improvements of the International Reference Ionosphere model for the topside electron density profile', *Radio Science*, 41(RS5S15): pp. 1–8, 2006.

Bilitza, D. and Williamson, R., 'Towards a better representation of the IRI topside based on ISIS and Alouette data', *Advances in Space Research*, 25(1): pp. 149–152, 2000.

Blanch, J., *Using kriging to bound satellite ranging errors due to the ionosphere*, Ph.D. thesis, Stanford University, Stanford, California, USA, 2003.

Booker, H. G., 'A theory of scattering of nonisotropic irregularities with application to radar reflection from the aurora', *Journal of Atmospheric and Terestrial Physics*, 8: pp. 204–221, 1956.

Booker, H. G., 'Fitting of multi-region ionospheric profiles of electron density by a single analytic function of height', *Journal of Atmospheric and Terrestrial Physics*, 39: pp. 619–623, 1977.

Brace, L. H. and Theis, R. F., 'Global empirical models of ionospheric electron temperature in the upper F-region and plasmasphere based in in situ measurements from the Atmosphere Explorer-C, ISIS-1 and ISIS-2 satellites', *Journal of Atmospheric and Solar-Terestrial Physics*, 43(12): pp. 1317–1343, 1981.

Bradley, P. A. and Dick, M. I., 'Use of ground-based and satellite data for an improved procedure for testing the accuracy of ionospheric maps', Technical report, Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX, UK, 1997.

Calais, E. and Minister, J. B., 'GPS, earthquakes, the ionosphere and the space shuttle', *Physics Of The Earth and Planetary Interiors*, 105: pp. 167–181, 1998.

Cander, L. R., 'Ionospheric research and space weather services', *Journal of Atmospheric and Solar-Terestrial Physics*, 70: pp. 1870–1878, 2008.

Cander, L. R., Leitinger, R. and Levy, M. F., 'Ionospheric models including the auroral environment', *in proceedings of the ESA WPP-155*, ESA Workshop on Space Weather, ESTEC, Noordwijk, 1998.

Carlson, H. C. and Gordon, W. E., 'Radar spectrographic estimates of ionic composition from 225 -1400 km for solar minimum winter and summer conditions', *Journal of Geophysical Research*, 71(23): pp. 5573–5578, 1966.

Carpenter, D. L., 'Remote sensing the Earth's plasmasphere', *Radio Science Bulletin*, 308: pp. 13–29, 2004.

Chapman, S., 'The absorption and dissociative or ionising effect of monochromatic radiation in an atmosphere on a rotating earth', *in proceedings of the Physical Society*, volume 43, pp. 26–45, 1931.

Chapman, S., 'Geomagnetic nomenclature', *Journal of Geophysical Research*, 68: p. 1174, 1963.

Choi, B. K., Park, J. U. and Chung, J. K., 'Ionospheric tomography using a regional GPS network over South Korea', *Journal of Global Positioning Systems*, 5(1-2): pp. 47–51, 2006.

Coïsson, P., Radicella, S. M., Leitinger, R. and Nava, B., 'Topside electron density in IRI and NeQuick: features and limitations', *Advances in Space Research*, 37(5): pp. 937–942, 2006.

Coïsson, P., Radicella, S. M. and Nava, B., 'Comparisons of experimental topside electron concentration profiles with IRI and NeQuick models', *Annales Geophysicae*, 45(1): pp. 125–130, 2002.

Davies, K., *The upper atmosphere - data analysis and interpretation*, Ionosphere models, Springer-Verlag, New York, 1996.

Davies, K. and Hartmann, G. K., 'Studying the ionosphere with the Global Positioning System', *Radio Science*, 32(4): pp. 1695–1703, 1997.

Defrise, M., 'A short reader's guide to 3d tomographic reconstruction', *Computerised Medical Imaging and Graphics*, 25(2): pp. 113–116, 2001.

Depuev, V. H. and Pulinets, S. A., 'A global empirical model of the ionospheric topside electron density', *Advances in Space Research*, 34(9): pp. 2016–2020, 2004.

Di Giovani, G. and Radicella, S. M., 'An analytical model of the electron density profile in the ionosphere', *Advances in Space Research*, 10(11): pp. 27–30, 1990.

Dungey, J. W., 'The effect of ambipolar diffusion in the night time F layer', *Journal of Atmospheric and Terrestrial Physics*, 9: pp. 90–102, 1956.

Erturk, O., Arikan, O. and Arikan, F., 'Tomographic reconstruction of the ionospheric electron density as a function of space and time', *Advances in Space Research*, 43(11): pp. 1702–1710, 2009.

Ezquer, R. G., Jadur, C. A. and de Gonzalez, M. M., 'IRI-95 TEC predictions for the South American peak of the equator anomaly', *Advances in Space Research*, 22(6): pp. 811–814, 1998.

Faletič, R., *Tomographic reconstruction of shock layer flows*, Ph.D. thesis, Australian National University, Canberra, Australia, 2005.

Fausett, L., *Fundamentals of neural networks, architectures, algorithms, and applications*, Prentice Hall International Inc, Eaglewood cliff, New Jersey, 1994.

Ferraro, V. C. A., 'Diffusion of ions in the ionosphere', *Journal of Geophysical Research*, 50: pp. 215–222, 1945.

Florida, C. D., 'The development of a series of ionospheric satellites', *in proceedings of the IEEE*, volume 57, pp. 867–875, 1969.

García-Fernández, M., *Contributions to the 3D ionospheric sounding with GPS data*, Ph.D. thesis, Universitat Politècnica de Catalunya, Spain, 2004.

Gillies, R. G., *Modelling of transionospheric HF radio wave propagation for the ISIS II and ePOP satellites*, Master's thesis, University of Saskatchewan, Saskatchewan, Canada, 2006.

Habarulema, J. B., McKinnell, L. A. and Cilliers, P. J., 'Prediction of Global Positioning System total electron content using neural networks over South Africa', *Journal of Atmospheric and Solar-Terestrial Physics*, 69(15): pp. 1842–1850, 2007.

Habarulema, J. B., McKinnell, L. A. and Opperman, B. D. L., 'A recurrent neural network approach to quantitatively studying solar wind effects on TEC derived from GPS; preliminary results', *Annales Geophysicae*, 27: pp. 2111–2125, 2009.

Hansen, A., *Tomographic estimation of the ionosphere using terrestrial GPS sensors*, Ph.D. thesis, Stanford University, Stanford, California, USA, 2002.

Hargreaves, J. K., *The solar-terrestrial environment: an introduction to geospace - the science of the terrestrial upper atmosphere, ionosphere, and magnetosphere*, Cambridge Atmospheric and Space Science Series, Cambridge University Press, New York, 1992.

Hartz, T. R., 'Observation of the galactic radio emission between 1.5 and 10 MHz From the Alouette satellite', *Annales d'Astrophysique*, 27: pp. 823–830, 1964.

Havens, R. J., Friedman, H. and Hulburt, E. O., 'The ionospheric F2 region', *in Report of the Physical Society Conference*, The physics of the ionosphere, Cavedish Laboratory, Cambridge, UK, 1954.

Hedin, A. E., 'Extension of the MSIS thermosphere model into the middle and lower atmosphere', *Journal of Geophysical Research*, 96(A2): pp. 159–172, 1991.

Heelis, R. A., Hanson, W. B. and Bailey, G. J., 'Distributions of He^+ at middle and equatorial latitudes during solar maximum', *Journal of Geophysical Research*, 95: pp. 10313–10320, 1990.

Heise, S., Jakowski, N., Wehrenpfennig, A., Reigber, C. and Lühr, H., 'Sounding of the topside ionosphere/plasmasphere based on GPS measurements from CHAMP: Initial results', *Geophysical Research Letters*, 29(14), 1699, doi:10.1029/2002GL014738, 2003.

Hoffman, J. H., 'A mass spectrometric determination of the composition of the nighttime topside ionosphere', *Journal of Geophysical Research*, 71: pp. 1883–1888, 1967.

Huang, X. and Reinisch, B. W., 'Automatic calculation of electron density profiles from digital ionograms. 2. true height inversion of topside ionograms with the profile-fitting method', *Radio Science*, 17(4): pp. 837–844, 1982.

Huang, X. and Reinisch, B. W., 'Vertical electron profiles from the digisonde network', *Advances in Space Research*, 18(6): pp. 121–129, 1996.

Huang, X., Reinisch, B. W., Bilitza, D. and Benson, R. F., 'New data on the topside electron density distribution', Technical Report NAG5-8145, Goddard Space Flight Center, Greenbelt, Maryland, USA, 2001.

Huang, X., Reinisch, B. W., Bilitza, D. and Benson, R. F., 'Electron density profiles of the topside ionosphere', *Annals of Geophysics*, 45(1): pp. 125–130, 2002.

Huang, X. Q. and Reinisch, B. W., ' Vertical electron content from ionograms in real time', *Radio Science*, 36(2): pp. 335–342, 2001.

Hulburt, E. O., 'Ionisation in the upper atmosphere of the Earth', *Physical Review*, 31: pp. 1018–1037, 1928.

Hunsucker, R. D. and Hargreaves, J. K., 'The high-latitude ionosphere and its effects on radio propagation', Cambridge University press, 2003.

Jackson, J. E., 'The reduction of topside ionograms to electron-density profiles', *in proceedings of the IEEE*, volume 57, pp. 960–976, 1969.

Jakowski, N., Kutiev, I. S., Heise, S. and Wehrenpfennig, A., 'A topside ionosphere/plasmasphere model for operational applications', *in proceedings of the XXVII URSI General Assembly, Maastricht, The Netherlands*, pp. 2174–2177, 2002.

Jakowski, N., Leitinger, R. and Angling, M., 'Radio occultation techniques for probing the ionosphere', *Annals of Geophysics*, 47(2/3): pp. 1049–1066, 2004.

Jakowski, N., Schluter, S. and Shadron, E., 'Total electron content of ionosphere during the geomagnetic storm of January 1997', *Journal of Atmospheric and Solar-Terestrial Physics*, 61: pp. 299–307, 1999.

Jeffrey, R. A., Steven, J. F. and Liu, C. H., 'Ionospheric imaging using computerised tomography', *Radio Science*, 23(3): pp. 299–307, 1988.

Komjathy, A., *Global ionospheric total electron content mapping using the Global Positioning System*, Ph.D. thesis, Department of Geodesy and Geomatics, University of New Brunswick, Fredericton, New Brunswick, Canada, 1997.

Kunitake, M., *Inverse methods*, volume 68 of *Lecture Notes in Earth Sciences*, Springer Berlin / Heidelberg, 1996.

Kutiev, I. and Marinov, P., 'Topside sounder model of scale height and transition height characteristics of the ionosphere', *Advances in Space Research*, 39: pp. 759–766, 2007.

Kutiev, I., Marinov, P. and Serafimov, K., 'An approximation of the height of O^+ - H^+ transition level for use in the IRI', *Advances in Space Research*, 4(1): pp. 119–121, 1984.

Kutiev, I., Stankov, S. and Marinov, P., 'Analytical expression of O^+-H^+ ion transition surface for use in IRI', *Advances in Space Research*, 14(12): pp. 135–139, 1994.

Kutiev, I. S., Marinov, P. G. and Watanabe, S., 'Model of topside ionosphere scale height based on topside sounder data', *Advances in Space Research*, 37(5): pp. 943–950, 2006.

Leitinger, R., 'A magnetic field aligned approach to model the topside F2 layer', *Advances in Space Research*, 22(6): pp. 789–792, 1998.

Leitinger, R., Radicella, S., Hochegger, G. and Nava, B., 'Diffusive equilibrium models for the height region above the F2 peak', *Advances in Space Research*, 29(6): pp. 809–814, 2002.

Liu, L., He, M., Wan, W. and Zhang, M. L., 'Topside ionospheric scale heights retrieved from constellation observing system for meteorology, ionosphere, and climate radio occultation measurements', *Journal of Geophysical Research*, 113(A10304), doi:10.1029/2008JA013490, 2008.

Liu, L., Le, H., Wan, W., Sulzer, M. P., Lei, J. and Zhang, M., 'An analysis of the scale heights in the lower topside ionosphere based on the Arecibo incoherent scatter radar measurements', *Journal of Geophysical Research*, 112(A06307), doi:10.1029/2007JA012250, 2007a.

Liu, L., Wan, W. and Ning, B., 'A study of the ionogram derived effective scale height around the ionospheric h_mF2', *Annales Geophysicae*, 24: pp. 851–860, 2006.

Liu, L., Wan, W., Zhang, M. L., Ning, B., Zhang, S. R. and Holt, J. M., 'Variations of topside ionospheric scale heights over Millstone Hill during the 30-day incoherent scatter radar experiment', *Annales Geophysicae*, 25: pp. 2019–2027, 2007b.

Liu, Z., Skone, S., Gao, Y. and Komjathy, A., 'Ionospheric modelling using GPS data', *GPS Solutions*, 9(1): pp. 63–66, 2005.

Luan, X., Liu, L., Wan, W., Lei, J., Zhang, S., Holt, J. M. and Sulzer, P. M., 'A study of the shape of the topside electron density profile derived from incoherent scatter radar measurements over Arecibo and Millstone Hill', *Radio Science*, 41(RS4006), doi:10.1029/2005RS003367, 2006.

Marinov, P., Kutiev, I. and Watanabe, S., 'Empirical model of O^+ - H^+ transition height based on topside sounder data', *Advances in Space Research*, 34(9): pp. 2015–2022, 2004.

McKinnell, L. A., *A neural network based ionospheric model for the bottomside electron density profile over Grahamstown, South Africa*, Ph.D. thesis, Rhodes University, Grahamstown, South Africa, 2002.

McKinnell, L. A. and Poole, A. W. V., 'Neural network-based ionospheric modelling over the South African region', *South African Journal of Science*, 100: pp. 519–523, 2004.

McNamara, L. F., *The ionosphere: Communications, surveillance and direction finding*, Kriegner publishing company, Malabar, Florida, 1991.

Mehrotra, K., Mohan, C. K. and Ranka, S., *Elements of artificial neural networks*, The MIT press, Cambridge, MA 02142, USA, 1996.

Miyazaki, S., 'Ion transition height distribution obtained with TAIYO', *Journal of Geomagnetism and Geoelectricity*, 31: pp. 113–124, 1979.

Munnucci, A. J., Tsurutani, B. T., Iijima, B. A., Komjathy, A., Saito, A., Gonzalez, W. D., Guarnieri, F. L., Kozyra, J. U. and Skoug, R., 'Dayside glonal ionospheric response to the major interplanetary events of October 29–30. 2003 "Halloween Storms"', *Geophysical Research Letters*, 32(L12S02), doi:10.1029/2004GL021467, 2005.

Muruyama, T., Ma, G. and Nakamura, N., 'Signature of TEC storm on 6 November 2001 derived from dense GPS receiver network and ionosonde chain over Japan', *Journal of Geophysical Research*, 109(A10302), doi:10.1029/2004JA010451, 2004.

Opperman, B. D. L., *Reconstructing ionospheric TEC over South Africa using signals from a regional GPS network*, Ph.D. thesis, Rhodes University, Grahamstown, South Africa, 2007.

Oyama, K. I., Balan, N. and Watanabe, S., 'Morning overshoot of Te enhanced by downward plasma drift in the equatorial topside ionosphere', *Journal of Geomagnetism and Geoelectricity*, 48: pp. 959–966, 1996.

Paradza, M. W., *Development of a neural network based model for predicting the occurrence of spread F for the Brazilian sector*, Master's thesis, Rhodes University, Grahamstown, South Africa, 2008.

Pulinets, S., Depuev, V., Karpachev, A., Radicella, S. and Danilkin, N., 'Recent advances in topside profile modelling', *Advances in Space Research*, 29(6): pp. 815–823, 2002.

Radicella, S. M. and Leitinger, R., 'The evolution of the DGR approach to model the electron density profiles', *Advances in Space Research*, 27(1): pp. 35–40, 2001.

Radicella, S. M. and Zhang, M. L., 'The improved DGR analytical model of the electron density height profile and total electron content in the ionosphere', *Annales Geophysicae*, 38(1): pp. 35–41, 1995.

Rama Rao, P. V. S., Gopi Krishna, S., Niranjan, K. and Prasad, D. S. V. V. D., 'Temporal and spatial variations in TEC using simultaneous measurements from the Indian GPS network of receivers during the low solar activity period of 2004–2005', *Annales Geophysicae*, 24: pp. 3279–3292, 2000.

Rama Rao, P. V. S., Gopi Krishna, S., Niranjan, K. and Prasad, D. S. V. V. D., 'Temporal and spatial variations in TEC using simultaneous measurements from the

Indian GPS network of receivers during the low solar activity period of 2004–2005', *Annales Geophysicae*, 24: pp. 3279–3292, 2006.

Rawer, K., Bilitza, D., Rmakrishnan, S. and Sheikh, N., 'Intentions and build-up of the International Ieference Ionosphere', *in Operational modelling of the aerospace propagation environment*, edited by H. T. Yura, volume 1, 1978.

Reinisch, B. W., 'Tenth international digisonde training seminar at UMass Lowell reviews state of real time mapping of the ionosphere', *IEEE Antennas and Propagation Magazine*, 45(5): pp. 110 – 117, 2004.

Reinisch, B. W., Haines, D. M., Sales, G. S., Benson, R. F., Green, J. L. and Taylor, W. W. L., 'Radio sounding in space: magnetosphere and topside ionosphere', *Journal of Atmospheric and Solar-Terestrial Physics*, 63: pp. 87–98, 2001.

Reinisch, B. W. and Huang, X. Q., 'Deducing topside profiles and total electron content from bottomside ionograms', *Advances in Space Research*, 27(1): pp. 23 – 30, 2001.

Reinisch, B. W., Huang, X. Q., Belehaki, A., Shi, J. K., Zhang, M. L. and Ilma, R., 'Modelling the IRI topside profile using scale heights from ground-based ionosonde measurements', *Advances in Space Research*, 34(9): pp. 2026 – 2031, 2004.

Reinisch, B. W., Nsumei, P., Huang, X. and Bilitza, D., 'Modelling the F2 topside and plasmasphere for IRI using IMAGE/RPI and ISIS data', *Advances in Space Research*, 39(5): pp. 731–738, 2007.

Richards, P. G., 'Seasonal and solar cycle variations of the ionospheric peak electron density; comparison of measurement and model', *Journal of Geophysical Research*, 106(A7): pp. 12803 – 12819, 2001.

Richards, P. G., Chang, T. and Comfort, R. H., 'On the causes of the annual variation in the plasmaspheric electron density', *Journal of Atmospheric and Solar-Terrestrial Physics*, 62: pp. 935 – 946, 2000.

Richards, P. G. and Torr, D. G., 'Seasonal, diurnal and solar cyclical variations of the limiting H^+ flux in the Earth's topside ionosphere', *Journal Geophysical Research*, 90(A6): pp. 5261–5268, 1985.

Rishbeth, H., 'On the theory of diffusion in the ionosphere', *Geophysical Journal*, 41: pp. 311–317, 1975.

Rishbeth, H. and Borron, D. W., 'Equilibrium electron distributions in the ionospheric F2-layer', *Journal of Atmospheric and Terestrial Physics*, 18: pp. 234–252, 1960.

Rishbeth, H. and Garriott, O. K., *Introduction to ionospheric physics*, volume 14 of *International Geophysics series*, Academic press, New York and London, 1969.

Ruzhin, Y. Y., Shagimuratov, I. I., Kunitsyn, V. E., Depueva, A. K. and Razinkov, O. G., 'GPS-based tomographic reconstruction of the ionosphere', *Advances in Space Research*, 21(3): pp. 521–524, 1998.

Schaer, S., *Mapping and predicting the Earth's ionosphere using the Global Positioning System*, Ph.D. thesis, Astronomical Institute, University of Berne, Berne, Switzerland, 1999.

Schlüter, S., Stolle, C., Jakowski, N. and Jacobi, C., 'Monitoring of 3-dimensional ionospheric electron density distributions based on GPS measurements', *in the First CHAMP Science Meeting*, GeoForschungsZentrum, Potsdam, Germany, 2002.

Schunk, R. W., 'A mathematical model of the middle and high latitude ionosphere', *Pure and Applied Geophysics*, 127(2-3): pp. 255–303, 1988.

Schunk, R. W. and Sojka, J. J., 'Approaches to ionospheric modelling, simulation and prediction', *Advances in Space Research*, 12(6): pp. 317–326, 1992.

Shuiskaya, F. K., Gal'Perin, I. I., Serov, A. A., Baranets, N. V. and Kushnerevskii, I. V., 'Resonant heating of the ionospheric plasma by powerful radio pulses aboard the Intercosmos-19 and Cosmos-1809 satellites', *Planetary Space Science*, 38(2): pp. 173–180, 1990.

Sibanda, P. and McKinnell, L. A., 'Evaluating the IRI topside model for the South African region: An overview of the modelling techniques', *Advances in Space Research*, 44: pp. 707–714, 2009a.

Sibanda, P. and McKinnell, L. A., 'The applicability of existing topside ionospheric models to the South African region', *South African Journal of Science*, 105: pp. 387–390, 2009b.

Stankov, M. S. and Muhtarov, P. Y., 'Reconstruction of the electron density profile from the total electron content using upper transition level and vertical incidence sounding measurements', *Comptes Rendus de l'Academie Bulgare des Sciences*, 54(9): pp. 45–48, 2001.

Stankov, S. M., Jakowski, N., Heise, S., Muhtarov, P., Kutiev, I. and Warnant, R., 'A new method for reconstruction of the vertical electron density distribution in the upper ionosphere and plasmasphere', *Journal of Geophysical Research*, 108(A5), 1164, doi:10.1029/2002JA009570., 2003.

Stankov, S. M., Warnant, R. and Jodogne, J. C., 'Operational model for real-time reconstruction of the electron density profile using GPS-TEC measurements', *in proceedings of the XXVII URSI General Assembly, Maastricht, The Netherlands*, pp. 1571–1574, 2002.

Suard, N. and Arbesser-Rastburg, B., 'Ionosphere and navigation - a brief overview', *Space Communications*, 20(1-2): pp. 3–5, 2005.

Sutton, E. and Na, H., 'High resolution ionospheric tomography through orthogonal decomposition', *in proceedings of the ICIP-94*, volume 2 of *IEEE International Conference*, pp. 148–152, 1994.

Sutton, E. and Na, H., 'A block iterative algorithm for tomographic reconstruction of ionospheric electron density', *International Journal of Imaging Systems and Technology*, 7(3): pp. 238–245, 1996.

Thidé, B., 'Nonlinear physics of the ionosphere and LOIS/LOFAR', *Plasma Physics and Controlled Fusion*, 49: pp. 103–107, 2007.

Titheridge, J. E., 'Determination of ionospheric electron content from the Faraday rotation of geostationery satellite signals', *Planetary Space Science*, 20: pp. 353–369, 1972.

Titheridge, J. E., 'Ion transition heights from topside electron density profiles', *Planetary Space Science*, 24: pp. 229–245, 1976.

Titheridge, J. E., 'Temperatures in the upper ionosphere and plasmasphere', *Journal of Geophysical Research*, 103(A2): pp. 2261–2277, 1998.

Triskova, L., Truhlik, V. and Smilauer, J., 'On possible improvements of outer ionosphere ion composition model in IRI', *Advances in Space Research*, 29(6): pp. 849–858, 2002.

Venkatraman, S., *Nighttime behaviour of the equatorial topside ionosphere*, Ph.D. thesis, University of Texas at Dallas, William B. Hanson Centre for Space Sciences , Richardson, TX 75080, USA, 1999.

Verronen, P. T., *Ionosphere-atmosphere interaction during solar proton events*, Ph.D. thesis, University of Helsinki, Helsinki, Finland, 2006.

Wakai, N. and Matuura, N., 'Operation and experimental results of the ionosphere sounding satellite ISS-b', *Acta Astronautica*, 7: pp. 999–1020, 1980.

Warnant, R., 'The increase of ionospheric activity as measured by GPS', *Earth Planets Space*, 52: pp. 1055–1060, 2000.

Webb, P. A. and Essex, E. A., 'An ionosphere-plasmasphere Global Electron Density model', *Physics and Chemistry of the Earth*, 25(4): pp. 301–306, 2000.

Wen, D., Yuan, Y. and Ou, J., 'Monitoring the three-dimensional ionospheric electron density distribution using GPS observations over China', *Journal of Earth System Science*, 116(3): pp. 235–244, 2007.

Williscroft, L. A. and Poole, A. W. V., 'Neural networks, foF2, sunspot number and magnetic activity', *Geophysical Research Letters*, 23(24): pp. 3659–3662, 1996.

Yang, W. Y., Cao, W., Chung, T. S. and Morris, *Applied numerical methods using MATLAB*, John Wiley and Sons, Inc., 2005.

Yizengaw, E., *Imaging the Ionosphere*, Ph.D. thesis, La Trobe University, Bundoora, Victoria, Australia, 2004.

Yizengaw, E. and Essex, E. A., 'Storm time seasonal variation of TEC on the southern hemisphere mid-Latitude regions using signals from GPS satellite', *in proceedings of the OIST-4*, pp. 221–224, Copenhagen, Denmark, 2002.

Yonezawa, T., 'On the influence of the electron-ion diffusion on the electron density and height of the nocturnal F2 layer', *Journal of Radio Research Laboratories*, 2: pp. 125–136, 1955.

Yonezawa, T., 'On the influence of electron-ion diffusion exerted upon the formation of the F2 layer', *Journal of Radio Research Laboratories*, 5: pp. 165–187, 1958.

Yoshimura, R., Maruyama, T. and Kawamura, S., 'Application of dynamical equilibrium model to topside ionosphere specification', *in proceedings of the XXV111 URSI General Assembly*, New Delhi, India, 2005.