A contribution to TEC modelling over Southern Africa using GPS data

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

DOCTOR OF PHILOSOPHY

of

Rhodes University

by

John Bosco Habarulema

August 2010

Abstract

Modelling ionospheric total electron content (TEC) is an important area of interest for radio wave propagation, geodesy, surveying, the understanding of space weather dynamics and error correction in relation to Global Navigation Satellite Systems (GNNS) applications. With the utilisation of improved ionosonde technology coupled with the use of GNSS, the response of technological systems due to changes in the ionosphere during both quiet and disturbed conditions can be historically inferred. TEC values are usually derived from GNSS measurements using mathematically intensive algorithms. However, the techniques used to estimate these TEC values depend heavily on the availability of near-real time GNSS data, and therefore, are sometimes unable to generate complete datasets. This thesis investigated possibilities for the modelling of TEC values derived from the South African Global Positioning System (GPS) receiver network using linear regression methods and artificial neural networks (NNs). GPS TEC values were derived using the Adjusted Spherical Harmonic Analysis (ASHA) algorithm. Considering TEC and the factors that influence its variability as "dependent and independent variables" respectively, the capabilities of linear regression methods and NNs for TEC modelling were first investigated using a small dataset from two GPS receiver stations. NN and regression models were separately developed and used to reproduce TEC fluctuations at different stations not included in the models' development. For this purpose, TEC was modelled as a function of diurnal variation, seasonal variation, solar and magnetic activities. Comparative analysis showed that NN models provide predictions of GPS TEC that were an improvement on those predicted by the regression models developed. A separate study to empirically investigate the effects of solar wind on GPS TEC was carried out. Quantitative results indicated that solar wind does not have a significant influence on TEC variability. The final TEC simulation model developed makes use of the NN technique to find the relationship between historical TEC data variations and factors that are known to influence TEC variability (such as solar and magnetic activities, diurnal and seasonal variations and the geographical locations of the respective GPS stations) for the purposes of regional TEC modelling and mapping. The NN technique in conjunction with interpolation and extrapolation methods makes it possible to construct ionospheric TEC maps and to analyse the spatial and temporal TEC behaviour over Southern Africa. For independent validation, modelled TEC values were compared to ionosonde TEC and the International Reference Ionosphere (IRI) generated TEC values during both quiet and disturbed conditions. This thesis provides a comprehensive guide on the development of TEC models for predicting ionospheric variability over the South African region, and forms a significant contribution to ionospheric modelling efforts in Africa.

Publications from this thesis

Several papers have been published over the 3 year period during which the research for this thesis was undertaken, and the content of these papers is incorporated in this thesis. The references and relevant sections of this thesis where the work appears are shown below.

1. John Bosco Habarulema, Lee-Anne McKinnell, Pierre J. Cilliers and Ben D. L. Opperman (2009): Application of neural networks to South African GPS TEC modelling, Advances in Space Research 43 (11), 1711-1720. This paper aimed at selecting the latitudinal range between two GPS stations for dependent TEC modelling using NNs. The choice of distance separation between GPS receivers used in the final model was based mainly on this paper's results and its contents form the major part of chapter 4.

2. John Bosco Habarulema, Lee-Anne McKinnell, Ben D. L. Opperman (2009): A recurrent neural network approach to quantitatively studying solar wind effects on TEC derived from GPS; preliminary results, Annales Geophysicae 27 (5), 2111-2125. In an attempt to empirically investigate the link between solar wind variability and TEC, a full modelling approach was undertaken and the details of this paper form most of chapter 5.

3. John Bosco Habarulema, Lee-Anne McKinnell and Ben D. L. Opperman (2009): Towards a GPS-based TEC prediction model for Southern Africa with feed forward networks, Advances in Space Research; 44 (1), 82-92. This paper presented a comprehensive effort undertaken to develop a single TEC simulation model for South Africa and its results are presented in chapter 6.

4. John Bosco Habarulema, Lee-Anne McKinnell and Ben D. L. Opperman (2010): TEC measurements and modelling over Southern Africa during magnetic storms ; a comparative analysis; Journal of Atmospheric and Solar-Terrestrial Physics 72 (5-6), 509-520. This paper investigated the predictability and regional mapping of TEC during magnetic storms and the results and data analyses are presented in both chapters 6 and 7.

5. Lee-Anne McKinnell, John Bosco Habarulema, Pierre J Cilliers and Ben D. L. Opperman (2010): GPS assistance in modelling the Southern African ionosphere; Geoscience and Remote Sensing Symposium, 2009 IEEE International, IGARSS 2009, 883-885. This paper briefly described how the GPS data complements ionosonde measurements in ionospheric modelling over Southern Africa. Some examples of results in this short paper are included in chapter 6.

Acknowledgements

I would like to thank my supervisors (Drs. Lee-Anne McKinnell and Ben D. L. Opperman) for their guidance, valuable ideas and contributions towards the completion of my project. The continuous support in many aspects provided by Dr. Lee-Anne McKinnell is greatly appreciated. The Hermanus Magnetic Observatory (HMO) provided the working environment, facilities as well as some funding during my study for which I greatly appreciate. Thanks to the HMO community for the interactions and lively discussions during my time of studies.

Throughout my postgraduate study, I would like to acknowledge the South African National Astrophysics and Space Science Programme (NASSP), University of Cape Town for the continuous and consistent funding. Thanks to the South African National Research Foundation (NRF), International Union of Radio Science (URSI) and International Beacon Satellite Symposium 2010 (among other funding agencies) for funding my travel and subsistence to Prague, Czech Republic, Chicago, USA and Barcelona, Spain for the research visit, General assembly and conference respectively where results of this research were presented.

I would like to express my heartfelt thanks to the family of Mr and Mrs David Rwarinda for all around continuous support from my Secondary school to the present. Special thanks to my parents, brothers and sisters for their support and understanding while I embarked on "marathon" studying (from Primary to PhD level without taking a rest!). I am highly indebted to them.

The assistance of Patrick Sibanda, Daniel Okoh, Stefan Lotz and Nkanyiso Mbatha in computer related software and programming is highly acknowledged. My flatmates Mpho Tshisaphungo and Etienne Koen with whom I would have a lot of social discussions especially after long days' work played a big role in occupying my mind!. We had useful and enriching discussions which I greatly appreciate. The tennis games (which kept me from academics all the time) I played with Sibusiso Mthembu ("Bus") were very relaxing. To my office mates; Chris Ndiitwani and Jean Uwamahoro, your contributions in many ways were valuable. Thanks to many friends for their support in different ways. The space is not enough to mention all their names!

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Since the 1990's, Global Navigation Satellite Systems (GNSS) have offered a complementary, relatively effective and alternative method of characterising the temporal and spatial behaviour of the ionosphere (e.g. Mannucci et al., 1995; Komjathy, 1997; Schaer, 1999; Meggs, 2005). Due to its dispersive character, electromagnetic signals experience time delay (modulated codes) and advance (carrier phase) as they propagate through the ionosphere. This delay is directly proportional to the integral number of electrons in a unit cross-sectional area (usually referred to as total electron content, TEC) along the signal path extending from the satellite to the receiver on the ground, and inversely proportional to the square of the frequency of propagation (Hofmann-Wellenhof et al., 1992; Misra and Enge, 2006).

Prior to the GNSS era, TEC studies were mainly carried out with the aid of measurements obtained from geostationary satellites using radio beacon signals through the Faraday rotation techniques (e.g. Klobuchar, 1975; Sethia et al., 1978; Davies, 1980), and ionosonde measurements, which still act as validation/verification tools for algorithms developed to extract TEC from GNSS observations. Other instrumentation such as incoherent scatter radars were also used (e.g. Mathews, 1972) and have been utilised along with newly developed techniques (e.g. Makela et al., 2000). The American Global Positioning System (GPS) which transmits signals at mainly two L-band frequencies, L1 (1575.42 MHz) and L2 (1227.60 MHz), was the first GNSS constellation to be operational and is of significant use to ionospheric TEC studies both regionally and globally (Komjathy, 1997; Schaer, 1999; Yizengaw et al., 2005; Yilmaz et al., 2009, etc). As additional GNSS constellations such as Galileo (European GNSS), the Russian Global Navigation Satellite System (GLONASS) and the Chinese COMPASS Navigation System become operational, the spatial resolution will improve and become more beneficial to scientific research among other applications.

The research presented in this thesis uses TEC measurements (for 2000-2007) derived from GPS

observations to perform regional TEC modelling/mapping over Southern Africa. Ionosonde measurements have been used to primarily validate GPS derived data as well as the modelled TEC data. For TEC modelling, artificial neural networks and linear regression methods were investigated. The verification of GPS derived and modelled TEC for a particular period is simply carried out by doing a comparative analysis with the corresponding ionosonde data set. This is primarily because the ionosonde is limited to monitoring the ionosphere directly above its geographical location. Additionally, since the ionosonde provides measurements for the bottomside ionosphere, the expected results of the comparative analysis should reflect that TEC derived from the ionosonde is less than TEC derived from GPS as well as produced by a model designed to reproduce TEC taking into account higher altitudes in its input parameters. There are four ionosondes in South Africa compared to more than 40 GPS receiver stations with two locations where an ionosonde and a GPS receiver are co-located (at Grahamstown and Hermanus). The available data has limited this study to the utilisation of data from three ionosonde stations (including one co-located with a GPS receiver) and therefore at best data from three stations were used to validate the modelled TEC. The Hermanus ionosonde was installed in June 2008 and therefore its data does not fall within the time period covered in this study.

Ionospheric research over South Africa has been going on for the past five or more decades, especially the characterisation of the bottomside ionosphere with the aid of ionosonde measurements (Gledhill et al., 1947; Poole and Evans, 1985; Gledhill, 1991; Williscroft and Poole, 1996; Poole and Poole, 2002; McKinnell, 2002). Non-linear approximations such as the application of neural network techniques in ionospheric parameter predictions have also been tested on the South African ionosonde data (e.g. McKinnell, 2002; Oyeyemi et al., 2006). The recent installation of dual frequency GPS receivers expanded the South African ionospheric research to include: presentation of first results of electron density profiles obtained by tomographically reconstructing TEC from the South African GPS receiver network (Cilliers et al., 2004); constructing an algorithm to derive TEC from GPS observations (Opperman et al., 2007); validating GPS-based TEC with ionosonde TEC (McKinnell et al., 2007); GPS TEC predictions using the neural network technique for single site locations by considering multiple input parameters (Habarulema et al., 2009a,c) and mapping GPS-derived ionospheric TEC over South Africa for solar cycle (SC) 23 (Moeketsi et al., 2007). This thesis describes the efforts undertaken to build on the knowledge acquired from single station TEC modelling and to construct and develop a TEC simulation model over Southern Africa using the same non-linear technique and GPS data. The first results towards the development of a comprehensive GPS-based TEC prediction model that provides TEC variational patterns over any point in South Africa are presented in Habarulema et al. (2009b), and its subsequent statistical and comparative validation

by means of regional TEC mapping can be found in Habarulema et al. (2010).

1.1 Problem description and motivation

1.1.1 Problem description

The ionosphere is studied using different parameters, including TEC from GPS. The GPS receiver network in South Africa offers an opportunity for both real-time and post-processing applications. Real-time applications, such as navigation, require immediate GPS data processing and subsequent communication of products to the user community. The work described in this thesis relates to post-processing applications where data is archived and analysed later. The unpredictable behaviour of ionospheric parameters and the need to supplement ionosonde data creates a need to develop nowcasting and forecasting techniques. In addition, there are sometimes gaps in the database due to infrastructure failures or shutdown for maintainance purposes. This work offers an opportunity to recover lost data through predictions, so that the TEC variations at any time of the day(s) during the year(s) for the period and region where the model is valid (over any GPS receiver location in South Africa) can be historically inferred for post-processing purposes. Furthermore, not every location can be covered by GPS infrastructure due to logistics, thus this work also provides an option for approximating, reproducing and artificially generating TEC variations at any latitude and longitude within South Africa, based on the actual data-driven mechanism. This expands the spatial data coverage when investigating phenomena which occur as a result of ionospheric transient processes, such as magnetic storms. The ionosphere exhibits complex and unpredictable behaviour with some unexpected characteristics which result from the occurence of magnetic or ionospheric storms. This makes it very difficult to develop reliable mathematical algorithms based on the physics that takes into account all these variations, and therefore data-driven approaches that involve parallel processing of multiple non-linear and complex systems, e.g. artificial neural networks, have been found more desirable (e.g. Cander et al., 1998; Chan and Canon, 2002; Poole and Poole, 2002; Oyeyemi et al., 2006; Tulunay et al., 2006; Leandro and Santos, 2007; Senalp et al., 2008; Yilmaz et al., 2009, and some references therein). Thus this thesis describes the model developed to simulate TEC behaviour over Southern Africa using the artificial neural network technique.

1.1.2 Motivation

Understanding and broadening scientific knowledge around various space weather effects requires constant monitoring and study of the space weather processes and their causes, and design of mitigation procedures with the aim of reducing the potential dangerous consequences

on humans and technology-based activities. To this effect, regional and global data is required as it is well understood and known that ionospheric variability exhibits changes across the globe. Ionospheric data (both historic and current) over Africa is still scarce as a result of the sparse distribution of space weather instruments and logistical limitations, and this necessitates the development of improvisation techniques. Visualisation of ionospheric behaviour in at least two dimensions over a certain region at any time requires simulation models. Thus this work aims to use the available GPS data, along with the techniques that take into account interpolation and extrapolation methods, to simulate regional ionospheric TEC behaviour. These techniques fill the gaps in the available data due to among other things, infrastructure failures and the absence of GPS receivers in some areas. In 2007 the Hermanus Magnetic Observatory (HMO) was given a mandate by the International Space Environment Service (ISES) to become the regional warning centre for Africa. In the long term, it is hoped that simulated TEC maps over Africa will be among other space weather products and services provided to the relevant communities and organisations by the HMO Space Weather Warning Centre (HSWWC) and therefore one of the aims of this work was the development of processes by which this goal can be realised.

1.2 Specific project objectives

The work described in this thesis aims to address directly or indirectly the following specific objectives

- Establish/build a database of TEC over South Africa
- Identify parameters that influence TEC variability
- Investigate the predictability of TEC over Southern Africa during both quiet and disturbed conditions
- Utilise the local database to finally construct a GPS-based TEC simulation model for South Africa capable of reproducing TEC patterns for any relevant given set of geophysical parameters.

1.3 Document structure

This thesis comprises 8 chapters. In chapter 1 the problem description and motivation for this research are given. Chapter 2 describes the formation and the constituents of the ionosphere, and describes the neural network technique used for modelling TEC in this study. The application of artificial neural networks is a broad concept and therefore only the feed forward and

recurrent network techniques employed in this thesis are discussed. Chapter 3 provides details of the application of GPS to ionospheric studies, along with a brief description of the algorithm that was used to derive TEC data from South African GPS observations. Chapter 4 provides a feasibility study of the use of artificial neural network and regression analysis models developed at selected GPS receiver stations and used to predict TEC variability at different stations not included in the models' development (verification of the models with "unseen" data). Chapter 5 describes an attempt made to quantitatively study the relationship between the solar wind and TEC derived from GPS. Chapter 6 is an extension of chapter 4 and describes the development of the final GPS-based TEC simulation model over Southern Africa. Chapter 7 presents results and some applications of the developed South African TEC simulation model. Finally, chapter 8 provides the conclusions and makes suggestions for improving the modelling of TEC over South Africa and its expansion over the entire African continent.

Chapter 2

Background theory

This chapter presents theoretical background concerning ionospheric structure, relevant ionospheric characteristics and some techniques that have been employed for the modelling of different ionospheric parameters. The basics of ionospheric formation mechanisms as well as a description of ionospheric variability are provided. The neural network technique, which is the primary method used in this research to model ionospheric TEC, is briefly discussed with particular emphasis on feed forward and recurrent networks. The status of available ionospheric data and models, particularly focussing on South Africa's ionosphere, is also provided.

2.1 Introduction to the ionosphere

The ionosphere is a region of the Earth's atmosphere that lies between an altitude of $\sim$ 50 km and $\sim$ 1400 km. It owes its existence to the sun and is mainly formed by the interaction of extreme ultra violet (EUV) and solar X-ray radiations with the neutral atmosphere. Before the introduction of the term "ionosphere" by Watson-Watt in about 1926, terms such as "Heaviside layer" and "Appleton layer" were used (Rishbeth et al., 1996). The E and F terms denoted the electric field of signals reflected from the now known E and F layers respectively. Since then, the ionosphere has been a subject of intensive study for its different roles in radio wave propagation, the need to fully understand space weather dynamics and of recent, to develop techniques for dealing with ionospheric effects on GNSS and other technology-related applications (e.g. Hofmann-Wellenhof et al., 1992; Tulunay et al., 2006; Misra and Enge, 2006). It is a complex medium since it is coupled to other systems such as the solar wind and the magnetosphere. In fact the ionosphere forms the transition region between the neutral atmosphere and the fully ionized plasma of the magnetosphere (Baumjohann and Treumann, 1997). The ionosphere is categorised into lower, bottomside and topside ionosphere in the order of their existence with respect to altitude (upper altitudes for these regions are $\sim$ 90 km, 350 km and 1000 km respectively). The lower and bottomside ionosphere are further classified into different layers

as discussed in section 2.3. The topside ionosphere extends above the F2 peak of the F layer. Above the topside ionosphere is the plasmasphere which forms a boundary to the ionosphere and extends into outerspace. Fig. 2.1 shows a schematic illustration of the different ionospheric layers. Beyond the F2 layer, the ionospheric electron density reduces with increasing altitude.

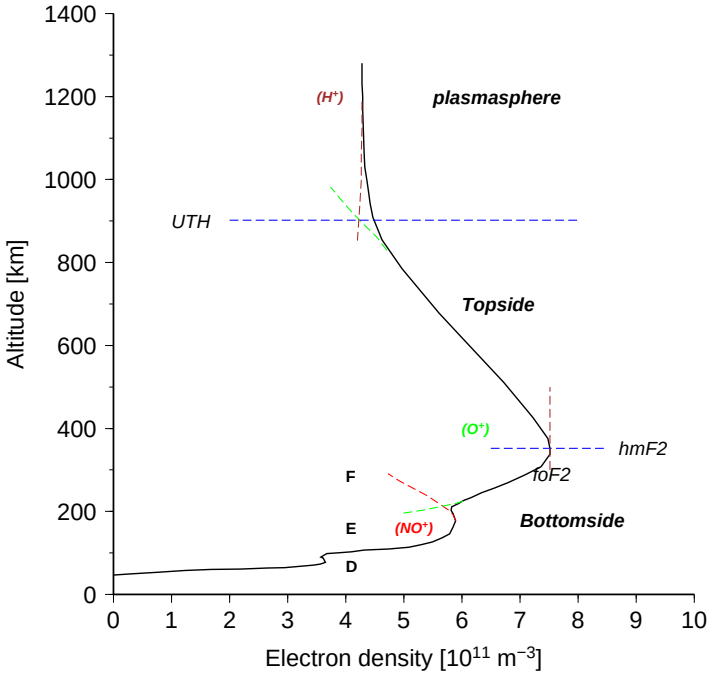


Figure 2.1: A schematic illustration of the ionospheric structure (Sibanda, 2010). UTH is the upper transition height, a point where the number of Oxygen ions and Hydrogen ions is approximately equal.

The topside ionosphere is characterised mainly by the presence of more H^+ ions compared to O^+ ions and subsequently H^+ ions are more dominant in the plasmasphere.

2.2 The composition of the ionosphere

The ionosphere can be considered a plasma and therefore consists of ions and neutral atoms. Ionisation is weak in the ionosphere in few relative extent between ions and neutral atoms. When these neutral atoms are hit by EUV light from the sun, electrons are stripped off thereby forming positively charged ions through a process known as photoionisation (McNamara, 1991; Baumjohann and Treumann, 1997), illustrated for the oxygen neutral atom as

$$O \longrightarrow O^+ + e^- \tag{2.1}$$

Photoionisation only occurs during the daytime when the sun is above the horizon and is greatest at about local midday (e.g. McNamara, 1991). Within the ionosphere, in absence of an

ionisation source, negatively charged electrons combine with positively charged ions to form neutral atoms. This process is called recombination and is literally the opposite of photoionisation (McNamara, 1991; Baumjohann and Treumann, 1997). Attachment, which is the process by which electrons combine or attach themselves to neutral atoms to form negatively charged ions, is the way electrons get lost in the lower levels (D region) of the ionosphere (e.g. McNamara, 1991). Both processes (attachment and recombination) always take place within the ionosphere (e.g. McNamara, 1991) and are negative contributors to ionisation (Baumjohann and Treumann, 1997). Recombination is further divided into two types: radiative and dissociative recombinations. Radiative recombination involves the direct combination of an electron with a positively charged ion to form a neutral atom and at the same time radiating some energy in the form of photon emission

$$O^+ + e^- \longrightarrow O + \text{photon (radiated energy)}. \tag{2.2}$$

On the other hand, dissociative recombination is a two stage process. The first process involves the interaction of positive ions with neutral molecules within the ionosphere, leading to the loss of one of the atoms (McNamara, 1991), e.g.

$$N_2 + O^+ \longrightarrow NO^+ + N. \tag{2.3}$$

The formed positively charged molecule combines with an electron to form two neutral atoms

$$NO^+ + e^- \longrightarrow N + O. \tag{2.4}$$

Radiative recombination is a slower process than dissociative recombination and therefore the latter leads to a greater loss of electrons than the former. More details about these mechanisms can be found in McNamara (1991).

2.3 Ionospheric layers

The ion composition, electron density and variability determine the altitude of different layers within the ionosphere. The most commonly known and studied layers are the D, E and F (F1 and F2) layers and they fall within the lower and bottomside ionosphere. There is also a minor C layer, believed to be produced by cosmic rays at an altitude of ~ 60 km, which some believe marks the boundary of the real ionosphere (e.g. Rishbeth et al., 1996). For each layer, there is a corresponding critical frequency (foD, foE, foF1 and foF2), which is the maximum frequency that can be reflected from that layer at vertical incidence (McNamara, 1991). Maximum electron density occurs within the F2 layer. The description of ionospheric layers in the following

subsections relies mainly on the information in Baumjohann and Treumann (1997).

2.3.1 D layer

This layer has an altitude range of $\sim$ 50 km to $\sim$ 90 km. This lower part of the ionosphere is weakly ionised since the recombination rate is high. This implies that it has low electron density even at its peak (in the order of $\sim 10^{10}$ m^{-3} at 90 km) compared to other layers. Due to the recombination process, the D layer tends to "dissappear" during the nighttime and its major source of ionisation is galactic cosmic rays. Ionisation is also generated by the X-rays emitted by solar flares which ionise gases within this layer leading to a strong absorption of radio waves (Ondoh and Marubashi, 2000).

2.3.2 E layer

This layer of the ionosphere falls within the altitude range of $\sim$ 90 km to $\sim$ 140 km. Its major source of ionisation is soft X-rays and UV solar radiation of molecular Oxygen. It is formed by the absorption of UV radiation (of long wavelength) that passes through higher layers until the molecular oxygen density rises to high levels below an altitude of $\sim$ 150 km (Baumjohann and Treumann, 1997). Mid-latitude electron densities at noon are of the order 10^{11} electrons per m^{-3}.

2.3.3 F layer

The F layer is divided into the F1 and F2 layers during the daytime. The F1 peaks at about 200 km and disappears during the night mainly due to the dissociative recombination process. The F2 layer reaches the peak at an altitude of $\sim$ 300 km (around local noon) and is always present throughout both daytime and nighttime hours. The F2 layer contains the maximum electron density (in the order of 10^6 cm^{-3}) within the ionosphere and the formation of the F layer is mainly due to the ionisation of oxygen atoms by solar EUV radiation.

2.4 Major ionospheric variations

The long and short term variations within the ionosphere have been monitored by means of the study of the temporal and spatial distribution of the electron density over an extended period of time. Most of the factors that describe the ionospheric variability are interlinked with the sun and its activities under different space weather conditions.

Solar cycle variations: The sunspot cycle is $\sim$ 11 years with periods of high and low solar activity variations. Since the sun is a major driver of space weather dynamics, high solar

activity directly translates to high ionospheric variability due into increased radiation that leads to an increase in photoionisation levels. Similarly, the ionosphere exhibits low variabilities during low solar activity levels.

Diurnal variation: The ionosphere's diurnal variation simply refers to day-night electron density variations as a result of the Earth's rotation with respect to the sun. Maximum electron density occurs during the day and is low for the night time period since there is no photoionisation as a result of the absence of solar radiation. In general, when photoionisation ceases, recombination greatly reduces the electron density during the night, but some free electrons still remain until dawn (McNamara, 1991).

Seasonal variations: The explanation of the seasonal variation of the ionosphere depends on the hemisphere over which the sun is overhead. However in both hemispheres, during summer the electron densities in the ionospheric layers can be expected to be higher than in winter, although this may not necessarily be the case, since the neutral atmosphere's composition also exhibits seasonal variations (e.g. McNamara, 1991; Meggs, 2005). At equinoxes (in March and September) the sun is directly overhead the equator during both solar maximum and minimum and this leads to increased electron density compared to other seasons.

Geomagnetic activity effects: Ultraviolet radiation from the sun is the source of thermal convection at ionospheric heights and this leads to the movement of ions and electrons across the geomagnetic field. The generated ionospheric current gives rise to a magnetic field in the ionosphere's surroundings, the variations of which are later observed as geomagnetic field fluctuations on the Earth's surface (Ondoh and Marubashi, 2000). During periods of geomagnetic storms, the ionospheric electron density is characterised mainly by negative and positive storm effects, depending among others on the latitudes and the strength of the disturbances.

Latitudinal variations: Over the equator the ionospheric densities are expected to be higher compared to other geographic locations. As the latitude increases on either side of the geographic equator, the solar radiation arrives in the atmosphere at an oblique angle, resulting in the reduction of ionospheric densities. The solar zenith angle (Z_A) determines the ionisation levels, thus geographic locations with low Z_A are exposed to more solar radiation and hence higher electron densities.

2.5 Ionospheric data sources and models in South Africa

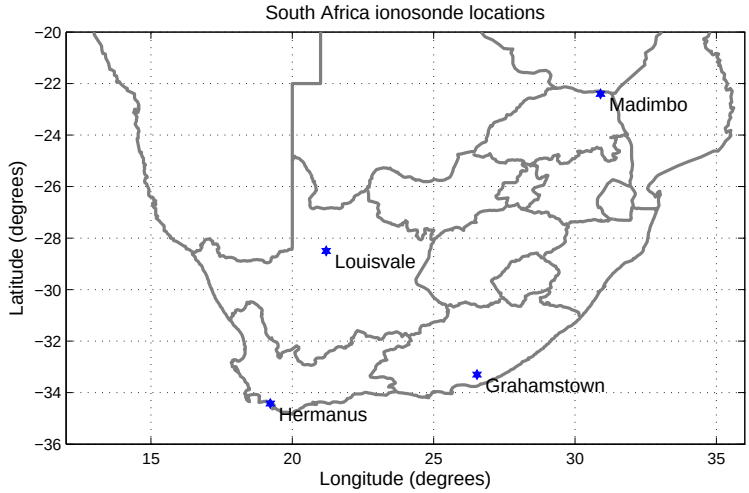
The conventional (and most common) way of recording ionospheric measurements was by means of ionosondes. Presently, GNSS receivers also provide ionospheric information (as discussed in chapter 3) over a relatively large area and have advantages of being portable and affordable. An ionosonde is a vertical sounding radar which measures ionospheric properties by vertical transmission and reception of radio wave signals at frequencies in the range $\sim$ 1-30 MHz. The ionosonde records measurements in the form of a graph of frequency against virtual height (h'), commonly known as the ionogram, where h' is the apparent height of the ionospheric layer derived from the time interval between the transmission and reception of signals. In other words, h' is the height at which a radiowave would be reflected had it continued to travel at the speed of light after entering the ionosphere. There are four operational ionosondes within South Africa; Grahamstown (33.3°S, 26.5°E), Louisvale (28.5°S, 21.2°E), Madimbo (24.4°S, 30.9°E) and Hermanus (34.43°S, 19.23°E). The Hermanus ionosonde was installed in June 2008 and provides the long-desired uniform coverage of the South African ionosphere. Fig. 2.2 shows the South African map with locations of the ionosondes, and an example of a typical ionogram.
Using ionosonde data, electron density values can be calculated from the vertically reflected radio waves. The relationship between the electron density and plasma frequency is given (Chen, 1984; McNamara, 1991) by

$$\omega^2 = \frac{Ne^2}{\epsilon_o m} \tag{2.5}$$

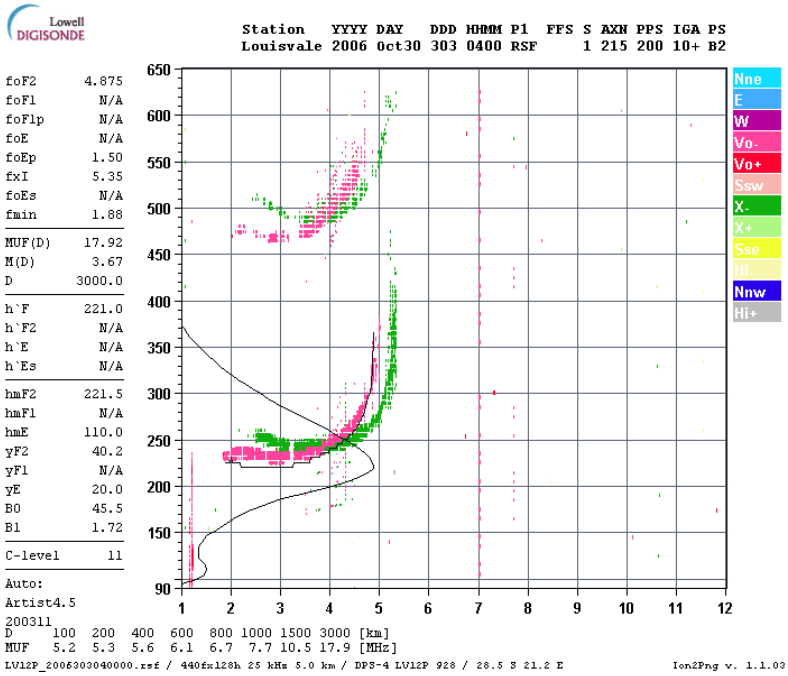
$$\omega = \frac{2\pi}{f_p} \tag{2.6}$$

where ω is the angular plasma frequency, N is the electron density (electrons per m^3), e is the electron charge ($\sim 1.60 \times 10^{-19}$ C), m is the electron mass ($\sim 9.11 \times 10^{-31}$ kg), ϵ_o is the permitivity of free space (8.854×10^{-12} Fm^{-1}) and f_p is the plasma frequency in Hz.

In South Africa, for many years, the ionospheric modelling focus has been on the characterisation of the bottomside ionosphere with the aid of ionosonde measurements (e.g. Gledhill et al., 1947; Poole and Evans, 1985; Gledhill, 1991; Williscroft and Poole, 1996; McKinnell, 2002; Poole and Poole, 2002; Oyeyemi et al., 2006). For more than a decade, the South African ionospheric researchers have concentrated mainly on artificial neural network modelling of ionospheric parameters for given geophysical conditions (e.g. Poole and McKinnell, 2000; McKinnell and Poole, 2003, 2004). A relatively new method to monitor and deduce ionospheric dynamics and variations has been explored with the use of GNSS observations from the South African GPS receiver network as well as the international GNSS receivers situated in the Southern Africa region (Opperman et al., 2007; Moeketsi et al., 2007). The ability to derive ionospheric



(a) Geographic location of South Africa's ionosondes.



(b) An example of ionogram from Louisvale ionosonde.

Figure 2.2: South Africa's ionosonde infrastructure and an example of an ionogram recorded at the Louisvale ionosonde station.

information from the South African GPS receiver network very recently expanded the studies to include the use of GPS-derived TEC along with the existing bottomside ionospheric information in efforts towards topside modelling over South Africa (Sibanda, 2010). The GPS measurements are particularly useful because they complement ionosonde data for plasmaspheric electron content characterisation. Therefore the present data sources mostly used for the ionospheric study in South Africa are ionosondes and the GPS. It is however important to mention that ionospheric dynamics can also be monitored using other sources such as incoher-

ent and coherent scatter radars, insitu instruments, etc. In this thesis, briefly decribed are the two ionospheric models (SABIM and IRI) for predicting South African ionospheric information, and one algorithm (ASHA) for deriving information from GPS data.

2.5.1 The South African Bottomside Ionospheric Model (SABIM)

SABIM was developed using bottomside ionospheric data from three South African ionosondes located at Grahamstown (33.3°S, 26.5°E), Louisvale (28.5°S, 21.2°E) and Madimbo (24.4°S, 30.9°E). It models the bottomside electron density profile over South Africa (McKinnell, 2002). SABIM is valid for the bottomside ionosphere (a region below $\sim$ 350 km in the atmosphere) and can, therefore, predict the electron density profile up to the peak height. It is a neural network-based model which was developed using input parameters (such as diurnal and seasonal inputs as well as solar and magnetic activity representations) that influence ionospheric variability.

2.5.2 The Adjusted Spherical Harmonic Analysis (ASHA) algorithm

The ASHA algorithm estimates TEC from GPS observations (Opperman et al., 2007). It is an analytical algorithm that relies on the availability of real and near real time data from the network of GPS receivers distributed throughout the country. By means of spherical harmonic interpolation one can use the ASHA algorithm to estimate TEC over any point within South Africa and generate two-dimensional TEC maps. The ASHA algorithm estimates TEC and eliminates the differential clock biases (DCBs). It was adapted from the Schaer (1999) global model to be used as a regional model using data from a local GPS receiver network and was chosen to estimate single station TEC results for comparative purposes with ionosonde measurements. One of the limitations of this algorithm is that the success of the TEC derivation heavily depends on the availability of GPS measurements. Data gaps exist due to the fact that GPS receivers may experience interruptions in their operations related, but not limited to, power problems, maintenance and other technical failures. It will be shown that through modelling/predictions, this limitation can be addressed (to ensure complete datasets) using the newly developed model. The TEC data used to develop the TEC simulation model for this thesis was based on this algorithm, the details of which are presented in chapter 3.

2.5.3 The International Reference Ionosphere (IRI) model

One of the most commonly used global models is the IRI. It describes the average ionospheric electron density, electron, neutral and ion temperatures as well as the ion composition as functions of height, location, local or universal time and sunspot number for magnetically quiet conditions (Bilitza, 2001). It is an international project sponsored by the International Union of Radio Science (URSI) and the Committee on Space Research (COSPAR). The IRI

is a standard empirical model that was developed based on all the available data sources (ionosondes, incoherent scatter radars, in-situ measurements on satellites and rockets, etc) and is updated periodically. TEC from the IRI is obtained by numerical integration in 1 km steps from 50 km to 2000 km. The IRI currently contains a geomagnetic activity dependence option based on the Time Empirical Ionospheric Correction Model (STORM) (Araujo-Pradere et al., 2002; Araujo-Pradere and Fuller-Rowell, 2002; Bilitza, 2003) which was formulated based on a study of the consistent and repeatable storm-time ionospheric response characteristics. The STORM model design includes seasonal dependence in the migration of the composition bulge by the global wind field and a non-linear dependence on the integrated time history of the geomagnetic activity index Ap. However, due to the general paucity of data from the Southern Hemisphere, the IRI does not give accurate ionospheric representation over Southern Africa in general, an observation that has previously been made based on bottomside ionospheric measurements as provided by the ionosonde data (McKinnell, 2002). However, in some cases, the IRI tends to give better TEC estimates than the regionally developed models for selected South African locations as will be shown in chapter 4.

2.6 Artificial neural networks (ANNs)

The development of ANNs was motivated by the desire to produce digital mechanisms that function similarly to the human brain. Just as the human brain adapts itself to new situations through learning, so does the artificial neural network (ANN). This means that artificial "neurons" are the information processing sytems in ANNs as neurons are to the human brain. Conceptually speaking, an ANN is computer software that has the ability to adapt information provided to it by learning its trends and patterns, storing the relevant information and generalising behaviour, using its parallelly distributed structure (Haykin, 1994). This section provides a theoretical description of the neural network (NN) technique with regard to non-linear approximation modelling. It is important to point out that it doesn't give complete details about this technique and the reader is referred to the references within this section for detailed descriptions. There are several algorithms available for ANN applications (Müller and Reinhardt, 1990; Watson, 1991; Haykin, 1994; Bishop, 1995; Fausett, 1994; Zell et al., 1998). However, the two generalised forms of ANNs explored in this study namely feed forward and recurrent networks are briefly discussed here.

2.6.1 Feed forward networks

Broadly, neural networks (NNs) are interconnected groups of "neurons" that are used to study the behavioural patterns of non-linear parameters through the storage of information in the interconnections and later generalising the variations of the parameter(s) under consideration

(Watson, 1991; Haykin, 1994). A NN architecture or setup consists of input, hidden and output layers. These layers can be customized to have a different number of neurons during the NN implementation. The number of neurons in the hidden layers depends on the amount of data being dealt with. For easy computation and the reduction in training time required for optimisation, the number of hidden layers is also subject to change (Haykin, 1994). Feed forward networks are a form of NNs, the architecture or topology of which does not permit closed paths within the entire network connections (Müller and Reinhardt, 1990; Haykin, 1994). Fig. 2.3 shows a schematic illustration of a feed forward network with two input neurons, n hidden neurons (h_1, h_2, ..., h_n) and one output neuron.

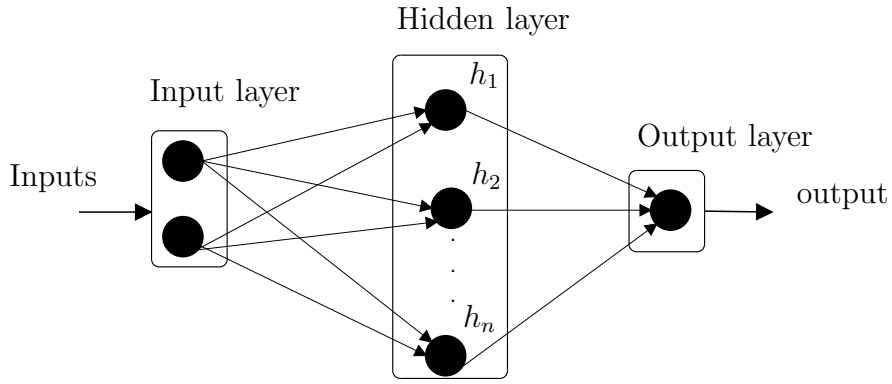


Figure 2.3: A schematic illustration of a simple feed forward network

2.6.1.1 Basic definitions for NN terminology

Before proceeding, the basic definitions for the terminology that will be consistently used are given below according to various sources (e.g. Müller and Reinhardt, 1990; Watson, 1991; Haykin, 1994; Bishop, 1995; Fausett, 1994).

1. Layers: These are sets of grouped neurons organised in a logical way, thus a layer is a single "block" consisting of small "blocks" or neurons.

2. Neurons: A neuron is an important element in a NN. It is just a simple processor that is connected to other neurons and does the function of summing up the output of a neuron or many neurons within the network, uses a specified mapping function to the current result and outputs the mapped sum to other inputs of other neurons (Watson, 1991). In a fully connected network, the output of neurons from the input layer becomes the input for neurons within the hidden layer(s) which later becomes the input for the neuron(s) in the output layer(s). This makes the neuron a fundamental unit in a NN. For data application purposes, the number of neurons is among the major determinants of generalisation and optimisation. A single neuron has the capacity to handle many inputs at once. Fig. 2.4

shows an example of inputs fed to the neuron and its output expressed as:

$$u_j = \sum_{i=1}^{k} w_{ji} x_i$$
$$y_j = \phi(u_j - \theta_j)$$

The external applied threshold (ϕ_j) lowers the effect of net input of the activation function, while the bias increases the net input of the function, thus the threshold is the negative of the bias (Haykin, 1994).

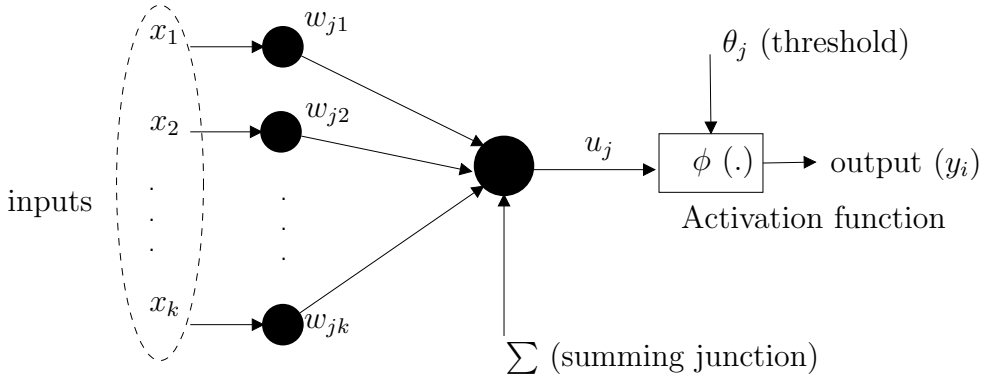


Figure 2.4: An example of the non-linearity of a neuron (Haykin, 1994).

3. Weights: A weight is a numerical determinant of the fraction of the neuron's output value that proceeds to the next neuron to which it is connected. The weights are randomly chosen at the start of the training in order to ensure generalisation.

4. Activation function: The mapping of input-output behaviour is performed by this function. For each neuron, the activation function determines the state of the neuron as a function of its bias/threshold of the weights of its incoming connections and of the states of the neurons connected to it by these connections (Müller and Reinhardt, 1990). The output of a neuron is known as its activation energy, implying that in the whole network, there are various activation energies involved (Watson, 1991).

In this study, feed forward networks were employed on the principle of supervised learning, which originally involves both known input training patterns and the desired system output patterns for the overall network during training (e.g. Watson, 1991; Haykin, 1994; Fausett, 1994; Bishop, 1995), so that the estimated difference between the output and target value is used as feedback to correct the output of the training pattern.

2.6.1.2 General feed forward algorithm

The feed forward network was simply implemented in topological order, using randomized weights with standard back propagation algorithm. These three features form the basis for the description of the following algorithm. As will be seen in chapters 4, 5 and 6, this study uses both single and two hidden layered networks, although the following description of an algorithm is for a network with one hidden layer. For the neuron input-output mapping processes, a binary sigmoid function, which is the default neuron activation function for the Stuttgart Neural Network Simulator (SNNS) software was used, i.e.

$$f(x) = \frac{1}{1 + e^{-ax}} \tag{2.7}$$

where $f(x)$ is the binary sigmoid function, $a = 1$ is the slope parameter and x is the input training vector. The above function is both differentiable, i.e $f'(x) = f(x)(1 - f(x))$ and continuous. These are among the key factors to consider when choosing activation functions for a back propagation algorithm. This algorithm has three main stages (Fausett, 1994), namely

- Directional "feeding" or feed forward of the input training vector to the input layer.
- The computation and back propagation of the errors.
- Adjustment of the randomly assigned synaptic weights.

The mathematical formulation and symbols in the following algorithm rely heavily on the material found in Fausett (1994).

1. Weight (w) initialisation: This is done randomly, falling within the range $-1 \leq w \leq 1$.
2. Application of feed forward:
   (i) Each input neuron (X_i, $i = 1, ..., n$) receives an input training pattern and forwards this pattern to all neurons in the hidden layer.
   (ii) Each hidden neuron (Z_j, $j = 1, ..., p$) sums its weighted input patterns

   $$z_j^{in} = v_{oj} + \sum_{i=1}^{n} x_i v_{ij} \tag{2.8}$$

   where v_{oj} is the bias on the hidden neuron j and v_{ij} is the connection weight between input neuron i and hidden neuron j.
   After the above step, the hidden neuron calculates its input pattern using its activation function $z_j = f(z_j^{in})$ and forwards it to all neurons in the output layer.

(iii) Each output neuron (Y_k, $k = 1, ..., m$) sums its weighted input patterns

$$y_k^{in} = w_{ok} + \sum_{j=1}^{p} z_j w_{jk} \tag{2.9}$$

where w_{ok} is the bias on the output neuron k and w_{jk} is the connection weight between the hidden neuron j and output neuron k.
It then applies its activation function to calculate its output pattern given by $y_k = f(y_k^{in})$. This marks the end of the feed forward process of the training patterns to the network.

3. Back propagation of errors:
(i) Each output neuron (Y_k, $k = 1, ..., m$) receives a target pattern corresponding to the input training pattern and calculates its error information term;

$$\delta_k = (t_k - y_k) f'(y_k^{in}), \tag{2.10}$$

computes the weight correction term $\Delta w_{jk} = \alpha \delta_k z_j$, proceeds to calculate the bias correction term $\Delta w_{ok} = \alpha \delta_k$, and finally sends δ_k to neurons in the hidden layer, where α is the learning parameter.
(ii) Each hidden neuron (Z_j, $i = 1, ..., p$) does the summation of its δ_k inputs as follows

$$\delta_j^{in} = \sum_{k=1}^{m} \delta_k w_{jk} \tag{2.11}$$

The error information term of each hidden neuron in this case is the product of the summation of the δ_k inputs and the derivative of its activation function, i.e $\delta_j = \delta_j^{in} f'(z_j^{in})$. After this, the weight correction term that is used to update v_{ij} is calculated as $\Delta v_{ij} = \alpha \delta_j x_i$, and finally the bias correction term becomes $\Delta v_{oj} = \alpha \delta_j$.

4. Weights and biases update:
(i) Each output neuron (Y_k, $k = 1, ..., m$) updates its weights and bias ($j = 0, ..., p$) as follows:

$$w_{jk}(\text{current}) = w_{jk}(\text{previous}) + \Delta w_{jk}. \tag{2.12}$$

(ii) Each hidden neuron (Z_j, $i = 1, ..., p$) also updates its weights and bias ($i = 0, ..., n$) as follows:

$$v_{ij}(\text{current}) = v_{ij}(\text{previous}) + \Delta v_{ij}. \tag{2.13}$$

The above process can't be implemented step by step as outlined, but is instead done iteratively for optimisation. The available data for NN implementation is divided into training, testing and

validation/verification patterns, with each pattern meant for a different task. As demonstrated in chapter 4, the training data set is the one used for training the NN through the adjustment of weights between the input, hidden and output layers, the validation/verification data set determines the performance of the NN on patterns that have not been trained, while the testing set is used to assess the final overall performance of the NN (e.g. Zell et al., 1998). With reference to the above algorithm, it is important to note that the weights update is done after the presentation of each training pattern, the very reason why a training algorithm is needed (Fausett, 1994).

2.6.2 Recurrent networks

These are a form of NNs with closed loop connections which are useful for learning time varying patterns (Fausett, 1994). Therefore, standard feed forward NNs are different from recurrent NNs in that the latter may involve outputs of hidden neurons being fed to themselves through the context or copy layer (e.g. Elman, 1990; Fausett, 1994; Haykin, 1994). This study uses one form of "partial" recurrent networks, known as Elman networks, that handles function learning by means of a combination of current input parameters and a series of previous learned parameters with the corresponding outputs of the entire network connection (Elman, 1990). Elman neural networks (ENNs) are considered partial recurrent networks because most of the connections are only feed forward (Fausett, 1994). This functional approach of ENNs makes possible the correction of the time delay between the various input parameters and the output parameter during the NN training and validating processes. In simple terms, an ENN is like a modified feed forward NN, the major difference being the presence of an additional layer consisting of context neurons which allows hidden unit patterns to be fed back to themselves (Elman, 1990). Since each hidden unit has a corresponding context unit in the ENN, the number of hidden neurons or nodes is equal to the number of context units in the context or copy layer (Marra and Marabito, 2005). Like other types of NNs, ENNs can have more than one hidden layer with the corresponding context layers. Other types of recurrent networks include Jordan network, Hopfield network, Echo state network, etc. Fig. 2.5 shows a simplified structure of the ENN comprising two input neurons, n hidden neurons and one output neuron. The outputs of each hidden neuron (h_1, h_2, ..., h_n) are propagated to each context neuron (c_1, c_2, ..., c_n) within the context layer and finally to the hidden layer again along with the incoming training pattern at time, $t+1$ from the input layer. This process continues iteratively until the number of iterations specified within the network is complete.

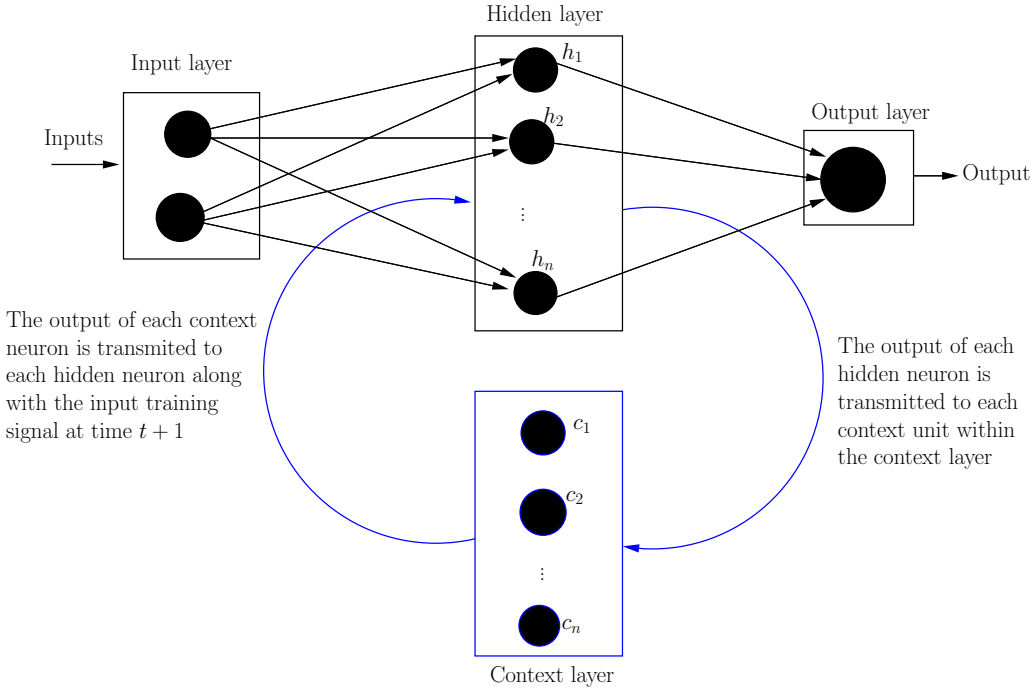


Figure 2.5: A schematic illustration of the ENN structure having one input, one hidden and one output layer.

2.6.2.1 A generalised and simplified algorithm

Consider a simple recurrent network consisting of one input layer, one hidden layer and one output layer. The presence of one hidden layer implies that there is a copy layer associated with it (Bodén, 2002; Marra and Marabito, 2005). The input training pattern forwarded through the weight layer is also combined with the previous state activation through an additional context layer, U. The following steps (details are found in Fausett, 1994; Bodén, 2002; Marra and Marabito, 2005) describe a simple recurrent network algorithm:

1. Recurrent network implementation:
   (i) The input training pattern or vector at time t is given by

$$y_j(t) = f(net_j(t)) \tag{2.14}$$

   For a recurrent network:

$$net_j(t) = \sum_{i}^{n} x_i(t)v_{ji} + \sum_{h}^{m} y_h(t-1)u_{jh} + \theta_j \tag{2.15}$$

   where n is the number of inputs, v_{ji} is the connection weight between input neuron i and hidden neuron j, u_{jh} is the connection weight between hidden neuron j and context neuron h and θ_j is the bias.

(ii) The network output is determined by the context layer and a set of output weights, w

$$y_k(t) = g(net_k(t)) \tag{2.16}$$

$$net_k(t) = \sum_{j}^{m} y_i(t) w_{kj} + \theta_k \tag{2.17}$$

g is the output function, m is the number of context neurons, θ_k is the bias, w_{kj} is the connection weight between hidden neuron j and output neuron k.

2. Back propagation for recurrent networks:
(i) Similar to feed forward networks, the sum square error (SSE) cost function is the most frequently used function (Bodén, 2002). Each training pattern p adds to this cost over all output neurons k (Fausett, 1994; Bodén, 2002).

$$C = \frac{1}{2} \sum_{p}^{n} \sum_{k}^{m} (d_{pk} - y_{pk})^2 \tag{2.18}$$

C is the cost function, d is the desired output, n is the total number of available training samples and m is the total number of output neurons.
(ii) Each weight change in the network should be proportional to the negative gradient of the cost function with respect to the specific weight that is subject to modification according to gradient descent (Bodén, 2002).

$$\Delta w = -\eta \frac{\partial C}{\partial w} \tag{2.19}$$

where η is the learning parameter. Using chain rule of differentiation, the weight change can be expressed in terms of error component $\delta = -\partial C/\partial net$ and $\partial net/\partial w$.
(iii) The errors for the output and hidden neurons respectively become

$$\delta_{pk} = -\frac{\partial C}{\partial y_{pk}} \frac{\partial y_{pk}}{\partial net_{pk}} = (d_{pk} - y_{pk}) g'(y_{pk}) \tag{2.20}$$

$$\delta_{pj} = -\left(\sum_{k}^{m} \frac{\partial C}{\partial y_{pk}} \frac{\partial y_{pk}}{\partial net_{pk}} \frac{\partial net_{pk}}{y_{pj}} \right) \frac{\partial y_{pj}}{\partial net_{pj}} = \sum_{k}^{m} \delta_{pk} w_{kj} f'(y_{pj}) \tag{2.21}$$

(iv) Assuming a first order polynomial, $\partial net/\partial w$ is simply the activation function, and

thus the weight change for the output and input weights become

$$\Delta w_{kj} = \eta \sum_{p}^{n} \delta_{pk} y_{pj} \tag{2.22}$$

$$\Delta w_{ji} = \eta \sum_{p}^{n} \delta_{pj} x_{pi}, \tag{2.23}$$

(v) Finally, the recurrent weights are modified according to the following equation after including the time component

$$\Delta u_{jh} = \eta \sum_{p}^{n} \delta_{pj}(t) y_{ph}(t-1). \tag{2.24}$$

The above decription of recurrent networks along with the notations for input, hidden (context) and output neurons is based on Bodén (2002), where clarification and details can be found.

2.6.3 Stuttgart Neural Network Simulator (SNNS)

SNNS is a project that was undertaken by computer scientists, including researchers from the universities of Stuttgart, Germany (Institute for Parallel and Distributed High Performance Systems (IPVR)) and Tübingen, Germany (Wilhelm Schickard Institute for Computer Science), in collaboration with the European Particle Research Lab, CERN, Geneva, Switzerland (Zell et al., 1998; Reczko et al., 1998). Their objective was to develop an efficient and portable neural network simulator. The development of the software along with some examples is detailed in Zell et al. (1998). SNNS offers flexibility and compatibility and is the major software used in this thesis for TEC prediction/modelling purposes. It was implemented within the Linux operating system and provides the relevant options for the two forms of ANNs (feed forward and recurrent networks) utilised in this study. Table 2.1 shows the chosen functions for feed foward and Elman networks in the order of their implementation in this thesis. The selected

Function	Feed forward networks	Elman networks
Update function	Topological order	Jordan Elman (JE) Order
Learning function	Standard Back Propagation	JE Back Propagation
Initialisation function	Randomised Weights	JE_Weights
Pattern remap function	None	None

Table 2.1: Functions for both feed forward and recurrent (Elman) networks (Zell et al., 1998)

functions have differences and similarities with regard to the two NN approaches and are briefly explained according to Zell et al. (1998) and Reczko et al. (1998) as follows:

- In the back propagation algorithm the learning parameter specifies the step width of the gradient descent, and the maximum difference (D_{max}) between the teaching value (t_j) and output value (o_j) that is propagated back is 0 i.e $D_{max} = t_j - o_j = 0$. In both the feed forward and Elman networks this applies only that in the latter an additional step of involving the context units exists.

- For the update function, topological order simply means that layers are processed in their order of arrangement according to the network topology (with the first hidden, second hidden, etc and finally the output) and thus neurons compute their activations topologically. JE_Order transmits a training pattern from the input layer to the hidden layer(s) and finally to the output layer. The update of all context units is done in a synchronous manner.

- For the initialisation function, the concept of randomised weights simply implies that all weights and biases are initialised in a random distribution within a certain range. While this is the only step in feed forward networks, JE_ Weights in Elman networks, involves at least four processes. Weights for forward connections are also randomly chosen. In addition, there is initialisation of weights of self-recurrent networks from context units to themselves, weights of other recurrent links to other context units and initialisation of all context units. Clearly, the context layer makes Elman networks significantly different from feed forward networks.

- A remapping function is a way of quickly varying the desired output of the network without changing pattern files. A remapping function of none as depicted in Table 2.1 for both feed forward and Elman networks, simply means that no remapping was done and thus all presented patterns were trained.

The SNNS software has a variety of inbuilt functions for optimisation and the reader is referred to Zell et al. (1998) for a comprehensive description of all the features along with the advantages and disadvantages for choosing certain functions.

2.7 Some examples of ANNs applied to ionospheric data

ANNs have been used in different forms to model and simulate ionospheric behavioural characteristics. The foF2 and TEC parameters have been extensively studied and modelled because of their effect on radio wave propagation, navigation and HF communication systems and other space weather related applications (Hernández-Pajares et al., 1997; Cander, 1998; Sarma and Madhu, 2005; Tulunay et al., 2006). The application of NNs to ionospheric modelling was motivated by the complex nature of the ionosphere (e.g. Chan and Canon, 2002; Tulunay et al.,

2006). Its coupling to the magnetosphere and the strong influence of very dynamic processes, such as atmospheric tides and neutral winds, along with its "out of average" variability during magnetic and ionospheric storms, makes it increasingly difficult to model successfully by means of mathematically-based algorithms that take into account the physics of each contribution (e.g. Cander, 1998; Tulunay et al., 2006).

To some extent, enough literature describing efforts to reproduce the foF2 variability (based on the actual data) at almost all latitudes exists (Williscroft and Poole, 1996; Cander, 1998; Poole and McKinnell, 2000; Sarma and Madhu, 2005; Oyeyemi et al., 2006) and statistical analyses show that ANNs perform relatively well in approximating the measured foF2 data as a function of relevant geophysical parameters. ANNs have also been applied to other ionospheric standard characteristics in the D, E and F layers (e.g. Cander, 1998; McKinnell, 2002; McKinnell and Poole, 2003, 2004). Cander (1998) presented a comprehensive study for modelling foF2 and TEC parameters as functions of multiple inputs that included hour, daily sunspot number and geomagnetic disturbance storm index, Dst, using time delay MLP NNs. Another study that utilised multiple inputs in foF2 modelling was done by Oyeyemi et al. (2006), in which a global model referred to as the near real time (NRTNN) model was constructed by including day number, hour, magnetic and solar activities, solar zenith angle, magnetic inclination and declination, as well as the angle of meridian as inputs. The historic foF2 data from stations around the globe was the desired output for the feed forward NN to learn its variational pattern. These studies demonstrate the capabilities of NNs in modelling ionospheric parameters by providing a history of the data physical processes that influence the ionosphere. While different studies may use NNs in different ways, for long term applications, diurnal and seasonal representations of the ionospheric variability should be taken into account when modelling any ionospheric parameter. Several other sources provide different methods of utilising NNs in modelling/predicting ionospheric parameters during different seasons, solar and magnetic activity periods, as well as at different latitudes (Poole and McKinnell, 2000; Poole and Poole, 2002; Chan and Canon, 2002; McKinnell and Poole, 2003; Sarma and Madhu, 2005; Oyeyemi et al., 2006; Leandro and Santos, 2007; Senalp et al., 2008; Tulunay et al., 2006; Yilmaz et al., 2009, and many others).

2.8 Summary

This chapter briefly discussed the mechanisms by which the ionosphere is formed and the availability of ionospheric data in South Africa. It was pointed out that of all the models that exist for South Africa, none provide predictions beyond 2000 km. Only the IRI provides TEC predictions up to 2000 km, but does not perform well over South Africa due to the data paucity for this region. The ASHA algorithm, which can provide TEC beyond 2000 km, relies heavily

on the availability of GPS observations. A brief description of the theoretical functioning of ANNs and their applications in non-linear approximation modelling was presented. Generalised algorithms for feed forward and recurrent networks were outlined with specific reference to Fausett (1994) and Bodén (2002) respectively. The software used for TEC predictions (SNNS) was introduced, with a focus on the four functions that mainly determine the performance efficiency and optimisation of a network. Only the two forms of NNs (feed forward and Elman networks) utilised in this study, were described. Finally, references to examples from the literature on NN data-driven models for foF2 and TEC variability were given.

Chapter 3

The GPS contribution to TEC studies

Ionospheric studies benefit from GPS data/measurements due to the fact that the ionosphere has a dispersive nature. Electromagnetic waves are refracted, reflected and also attenuated as they are transmitted through the ionised atmosphere. As stated in chapter 1, signals from the GPS satellites propagating at two L-band frequencies experience an unknown time delay (as they pass through the ionosphere) which to the first approximation, is directly proportional to TEC and inversely proportional to the square of the signal frequency. Thus, using this relation at two different frequencies, TEC, which can be considered a measure of ionisation in the ionosphere (e.g. Meggs, 2005), can be inferred from GPS measurements. This chapter gives a brief introduction to GPS, with emphasis on how TEC can be derived from GPS measurements. It also provides a description of the algorithm (Adjusted Spherical Harmonic Analysis, ASHA) that was used to estimate GPS TEC over South Africa, and upon which the regionally developed TEC simulation model was based. Finally, a brief description is given of the University of New Brunswick ionospheric mapping technique (UNB-IMT) algorithm (that was used for independent verification of the developed simulation model).

3.1 A brief introduction to the GPS

The GPS network is a constellation of at least 24 satellites inclined at an angle of 55° with respect to the equator. The network was established by the United States Department of Defense (USDoD), primarily for military purposes in order to provide accurate estimations of position, velocity and time. However, civilian use, including scientific research, surveying, precision agriculture, etc, was later allowed. The GPS network provides two main services; Standard Positioning Service (SPS) for civil use and the Precise Positioning Service (PPS) for the DoD and other authorised users. Originally, the signal meant for civil use was degraded by the introduction of controlled errors to reduce precision, a feature known as Selective Availability (S/A). This feature was later discontinued by a Presidential Decision Directive in May

2000 (Misra and Enge, 2006). However the USDoD can still limit the access to full capabilities of GPS through a mechanism known as Anti-Spoofing (AS). This is done by encrypting the P-code broadcast on both L1 and L2 to Y-code which can only be used authorised users (Misra and Enge, 2006). The GPS consists of three segments namely space, control and user segments. The system was designed and structured in such a way that at least four satellites are visible at or near the Earth at any one time. This is because at least four equations are required to solve the receiver's position and the time component from GPS observations. GPS satellites orbit at an altitude of about 20,000 km and transmit signals at two frequencies, f_1 (1575.42 MHz) and f_2 (1227.6 MHz). These signals propagate through different layers of the atmosphere and are attenuated especially in the ionosphere where the electron density is high. Further details, as well as a comprehensive literature survey on the development of the GPS network, can be found in Hofmann-Wellenhof et al. (1992) and Misra and Enge (2006).

3.1.1 GPS measurements

In the user segment domain, an observer has a choice of either using a single or dual frequency receiver in obtaining GPS measurements. Depending on the type of receiver used, different assumptions and mathematical computations are required to estimate ionospheric delay. Each GPS satellite transmits two Pseudo Random Noise (PRN) codes (Precise, P and Coarse Acquisition, C/A codes) and a navigation message for its identification and position determination respectively. Single frequency receivers track the C/A code only on the L1/f_1 frequency and mathematical algorithms incorporated into the receiver's hardware are needed to estimate ionospheric delay. Alternatively, the ionospheric delay for single frequency receivers can be estmated by applying an external correction to the position solution (e.g. Meggs, 2005). On the other hand, military dual frequency receivers track both PRN codes and the ionospheric delay can be estimated directly by comparing the delays on the L1 and L2 frequencies. However, civilian GPS receivers use special signal processing techniques to achieve both carrier phase and pseudo-range observables. Therefore, for the estimation of TEC, dual frequency receivers are preferred, since they are designed to eliminate ionospheric error which is also the major source of error in GPS measurements. Thus GPS measurements are classified as either code pseudo-ranges (P) or carrier phase (Φ) observables.

3.1.1.1 Pseudo-range measurements

A GPS receiver is used to determine the time that a signal takes to travel from a satellite to the receiver's position. This apparent transit time is defined as the difference between the signal reception time (as recorded by the receiver) and the satellite transmission time (Misra and Enge, 2006). However, the satellite and receiver clocks are independent of each other

(non-synchronous) and therefore the value of this transit time is biased by errors emanating from these clocks (e.g. Hofmann-Wellenhof et al., 1992; Misra and Enge, 2006). The product of this biased time and the speed of light in a vacuum gives the pseudo-range measurement. The pseudo-range measurements (P_1 and P_2) at L1 and L2 are given (e.g. Gao and Liu, 2002; Barrile et al., 2006) by

$$P_1 = \rho + c(dt - dT) + d\rho + d_T + I_1 + b^S_{P_1} - b^R_{P_1} + \varepsilon(P_1) \tag{3.1}$$

$$P_2 = \rho + c(dt - dT) + d\rho + d_T + I_2 + b^S_{P_2} - b^R_{P_2} + \varepsilon(P_2) \tag{3.2}$$

where

ρ: True geometric range (m) between the satellite and the receiver,
$d\rho$: Orbital error (m),
c: Speed of light (3.0×10^8 m/s),
dt: Satellite clock error with respect to GPS time (s),
dT: Receiver clock error with respect to GPS time (s),
d_T: Tropospheric error (m),
I_1, I_2: Pseudo-range ionospheric delays (m) at L1 and L2 respectively,
$b^s_{P_1}$, $b^s_{P_2}$: Pseudo-range satellite delays (m) at L1 and L2,
$b^r_{P_1}$, $b^r_{P_2}$: Pseudo-range receiver delays (m) at L1 and L2,
$\varepsilon(P_1)$, $\varepsilon(P_2)$: Pseudo-range measurement noises which include multipath errors (m).

If other errors (thermal noise and multipath) are ignored, differencing the pseudo-range measurements eliminates orbital error, geometric range, clock errors and tropospheric delay, and renders the "geometry-free" linear combination, i.e.

$$P_2 - P_1 = I + b^s_P + b^r_P \tag{3.3}$$

where

$$I = I_2 - I_1, b^s_P = b^s_{P_2} - b^s_{P_1}, b^r_P = b^r_{P_1} - b^r_{P_2}$$

By defining ionospheric delay as

$$I = \frac{40.28}{f^2}\text{TEC}, \tag{3.4}$$

where 40.28 is the constant of proportionality that can be obtained from the expression of the ionosphere's refractive index in subsection 3.1.2, the difference between the pseudorange measurements becomes

$$P_2 - P_1 = 40.28\left(\frac{1}{f_2^2} - \frac{1}{f_1^2}\right)\text{TEC} + b^s_P + b^r_P \tag{3.5}$$

And solving for TEC in terms of pseudorange measurements at the two frequencies yields

$$\text{TEC} = \frac{1}{40.28}\left(\frac{(f_1 f_2)^2}{f_1^2 - f_2^2}\right)\{(P_2 - P_1) - (b_P^s + b_P^r)\} \tag{3.6}$$

When the bias terms are ignored (in equation 3.6) and the respective values of f_1 and f_2 are substituted, TEC can be expressed in its units (TECU) as

$$\text{TEC} = 9.524(P_2 - P_1)\ \text{TECU} \tag{3.7}$$

where 1 TECU is equivalent to 10^{16} electrons.m^{-2}.

3.1.1.2 Carrier phase measurements

For a GPS receiver, the carrier phase measurement is the difference between the phases of the carrier signal generated by the receiver and the satellite carrier signal received at the time of measurement (e.g. Misra and Enge, 2006). This carrier phase is simply measured by determining the number of cycles received or generated since the starting point of the signal transmission. Since this involves assuming the distance between the satellite and receiver as a combination of an unknown number of whole cycles and a fraction of a cycle, this is an indirect and ambiguous measurement of the signal transit time (Misra and Enge, 2006). At L1 and L2 frequencies, the carrier-phase observation equations are given (e.g. Gao and Liu, 2002; Barrile et al., 2006) by

$$\Phi_1 = \rho + c(dt - dT) + d\rho + d_T + \lambda_1 N_1 - I_1 + b_{\Phi_1}^s - b_{\Phi_1}^r + \varepsilon(\Phi_1) \tag{3.8}$$

$$\Phi_2 = \rho + c(dt - dT) + d\rho + d_T + \lambda_2 N_2 - I_2 + b_{\Phi_2}^s - b_{\Phi_2}^r + \varepsilon(\Phi_2) \tag{3.9}$$

where

λ_1, λ_2: Carrier signal wavelengths at L1 and L2 frequencies respectively,
N_1, N_2: Carrier phase integer ambiguities,
I_1, I_2: Carrier phase ionospheric delays (m)
$b_{\Phi_1}^s$, $b_{\Phi_2}^s$: Carrier phase satellite delays or interfrequency biases (m)
$b_{\Phi_1}^r$, $b_{\Phi_2}^r$: Carrier phase receiver delays (m)
$\varepsilon(\Phi_1)$, $\varepsilon(\Phi_2)$: Carrier phase measurement noises with multipath errors included (m).

In particular, the carrier phase polarisation term corresponding to the product of the wavelength and the transmitter-receiver antenna relative rotating angle (which accounts for less than 0.5 TECU for ionospheric carrier phase combination L1-L2) is not included in eqns. 3.8 and 3.9.

Similar to the pseudo-range case, ignoring other errors, including multipath and thermal noise, the difference between the carrier-phase measurements eliminates the orbital error, geometric

range, clock errors and tropospheric delay, i.e.

$$\Phi_1 - \Phi_2 = \lambda N + I + b^s_\Phi + b^r_\Phi \tag{3.10}$$

where

$$\lambda N = \lambda_1 N_1 - \lambda_2 N_2, I = I_2 - I_1, b^s_\Phi = b^s_{\Phi_1} - b^s_{\Phi_2}, b^r_\Phi = b^r_{\Phi_2} - b^r_{\Phi_1}.$$

By definition the ionospheric phase advance is similar to the ionospheric delay (in magnitude) defined in equation 3.4. Using this result in equation 3.10 provides,

$$\Phi_1 - \Phi_2 = 40.28 \left(\frac{1}{f_2^2} - \frac{1}{f_1^2} \right) \text{TEC} + (\lambda_1 N_1 - \lambda_2 N_2) + (b^r_\Phi + b^s_\Phi) \tag{3.11}$$

Solving for a TEC value independent of ambiguities and biases gives

$$\text{TEC} = 9.524(\Phi_1 - \Phi_2) \ \text{TECU} \tag{3.12}$$

This is a precise but ambiguous measurement due to the presence of the integer ambiguities, while equation 3.7 provides an unambiguous but imprecise observable due to the presence of noise (Misra and Enge, 2006; Carrano and Groves, 2009). These two sets of observables are combined while using the pseudo-ranges to estimate the biases and ambiguities in the carrier-phase measurements.

3.1.2 Ionospheric effects on electromagnetic (EM) waves

The ionosphere's refractive index is complex and is the fundamental physical quantity required to understand the effects of EM waves as they traverse through the ionosphere. The refractive index reduces with increase in electron density (Ondoh and Marubashi, 2000), thus maximum attenuation of the signal is expected to occur within the ionosphere. In a dispersive medium like the ionosphere, the phase velocity from which the phase refractive index is derived, is a function of frequency (Langley, 2000). For system simplicity, consider a uniform external magnetic field **B** which makes an angle θ with the direction of propagation of a plane EM wave in the x-direction as shown in fig. 3.1. The expression for the complex refractive index of the ionosphere (commonly referred to as magnetoionic dispersion equation) was derived by Edward Appleton and Douglas Hartree in the early 1930's (Davies, 1990) as;

$$n^2 = 1 - \frac{X}{1 - jZ - \left(\frac{Y_T^2}{2(1-X-jZ)} \right) \pm \left(\frac{Y_T^4}{4(1-X-jZ)^2} + Y_L^2 \right)^{\frac{1}{2}}} \tag{3.13}$$

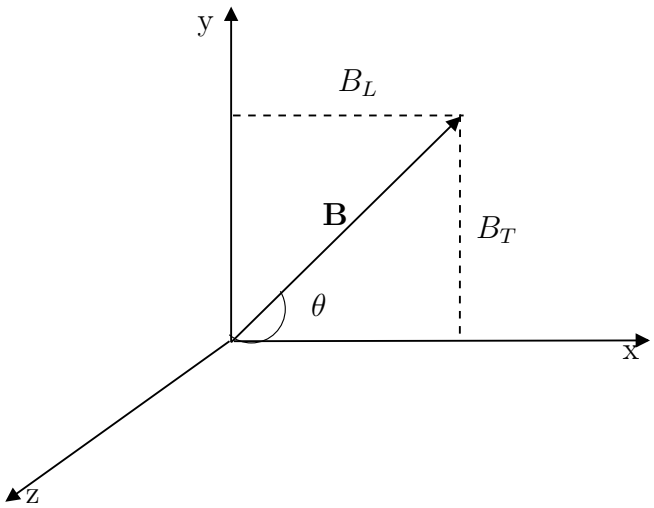


Figure 3.1: A schematic illustration of the system of orthogonal axes (after Davies, 1990)

where X, Y and Z are dimensionless quantities defined by the following expressions

$$X = \omega_p^2/\omega^2$$
$$Y = \omega_{\mathbf{B}}/\omega$$
$$Z = \omega_e/\omega$$

and

$$Y_L = \omega_L/\omega = \frac{\omega_{\mathbf{B}} cos\theta}{\omega}$$
$$Y_T = \omega_T/\omega = \frac{\omega_{\mathbf{B}} sin\theta}{\omega}$$

where

ω: Angular frequency of the propagating wave,
ω_p: Angular plasma frequency,
$\omega_{\mathbf{B}}$: Electron gyrofrequency, $\omega_{\mathbf{B}}$=**B**e/m,
ω_L: Longitudinal component of $\omega_{\mathbf{B}}$,
ω_T: Transverse component of $\omega_{\mathbf{B}}$,
ω_e: Angular collision frequency between electrons and heavier particles,
θ: Angle between the direction of propagation of the wave and the geomagnetic field,
B: Magnetic field strength.

For high frequencies, the collisional term may be ignored, i.e. $\omega_e \simeq 0$ and the Z term vanishes (Davies, 1990);

$$n^2 \cong 1 - \frac{X}{1 - \left(\frac{Y_T^2}{2(1-X)}\right) \pm \left(\frac{Y_T^4}{4(1-X)^2} + Y_L^2\right)^{\frac{1}{2}}} \tag{3.14}$$

Expanding equation 3.14 into the fourth inverse frequency power (e.g. Komjathy, 1997) gives

$$n \cong 1 - \frac{1}{2}X \pm \frac{1}{2}XY|\cos\theta| - \frac{1}{8}X^2 - \frac{1}{4}XY^2(1 + \cos^2\theta) \tag{3.15}$$

with the 2^{nd} and 4^{th} terms being functions of maximum electron density and frequency, while the 3^{rd} and 5^{th} are functions of electron density, the Earth's magnetic field strength and frequency (Komjathy, 1997). Komjathy (1997) further showed (through the use of typical values of the F2 layer electron density and the Earth magnetic field at L1 GPS frequency for $\theta = 0$) that the first and second terms of equation 3.15 can be used to estimate the refractive index, thereby neglecting other terms without introducing significant errors in its determination.
For EM wave propagation in the direction that is nearly perpendicular to $\mathbf{B}$ (quasi-transverse propagation), $\theta \simeq 90^\circ$ and $Y_L \simeq 0$ and equation 3.14 reduces to

$$n^2 \cong 1 - \frac{X}{1 - \left(\frac{Y_T^2}{2(1-X)}\right) \pm \left(\frac{Y_T^2}{2(1-X)}\right)} \tag{3.16}$$

The plane polarised EM wave will therefore be split into ordinary and extraordinary waves with the following expressions

$$n^2 \cong 1 - \frac{X(1-X)}{1 - X - Y_T^2} \qquad \text{extra-ordinary wave and}$$

$$n^2 \cong 1 - X \qquad \text{ordinary wave.}$$

The ordinary wave is the one that approximates the characteristics of a propagated wave without being affected by the magnetic field (e.g. Chen, 1984). Taking the refractive index n as complex ($n = \mu + i\chi$) and ignoring the effects of the magnetic field (Langley, 2000), both the ordinary and extraordinary wave expressions are the same:

$$n_\pm^2 \cong (\mu + i\chi)^2 = 1 - X. \tag{3.17}$$

Rewriting the real part (the refractive index must be real for the wave to propagate) in terms of the binomial expansion, the above equation gives

$$\mu^2 \cong 1 - X$$
$$\mu \cong 1 - \frac{1}{2}X + \frac{\frac{1}{2}(\frac{1}{2}-1)}{2!}X^2 + ...$$

Neglecting higher powers of X

$$\mu \approx 1 - \frac{1}{2}X = 1 - \frac{\omega_p^2}{2\omega^2}. \tag{3.18}$$

By means of the use of relevant assumptions, equation 3.18 can also be deduced from equation 3.15. The angular plasma frequency is given by

$$\omega_p^2 = \frac{N_e e^2}{m\epsilon_o}$$
$$\omega_p = 2\pi f_p, \quad \omega = 2\pi f$$

where

ω_p: Angular plasma frequency,
ω: Angular wave frequency,
N_e: Electron density (per m^3),
e: Charge of an electron (-1.602×10^{-19} C),
ϵ_o: Permittivity of free space (8.854×10^{-12} Fm^{-1}),
m: Electron rest mass (9.11×10^{-31} kg),
f_p: Plasma frequency (Hz),
f: Signal carrier frequency (Hz).

Thus

$$\mu \approx 1 - \frac{f_p^2}{2f^2} = 1 - \frac{N_e e^2}{8\pi^2 m\epsilon_o f^2}. \tag{3.19}$$

Substituting the respective values for electron mass, electron charge and the permittivity of free space, the above phase refractive index expression reduces to

$$\mu_\phi = 1 - \frac{40.28N_e}{f^2}. \tag{3.20}$$

The assumptions under which the Appleton-Hartree magneto-ionic theory is applicable include (e.g. Davies, 1990; Komjathy, 1997):

- The medium of EM wave propagation is electrically neutral, i.e. the number of ions and free electrons in the ionosphere are equal and there is no resultant space charge.

- Since the mass of an ion is greater than the electron mass, electrons are more mobile and thus positive ions have a negligible effect on radio waves.
- Ionospheric electrons oscillate at angular plasma frequency.

The group refractive index can be written in terms of the phase refractive index (Hofmann-Wellenhof et al., 1992; Langley, 2000) as

$$\mu_g = \mu_\phi + f\frac{d\mu_\phi}{df} \tag{3.21}$$

Differentiation of the phase refractive index gives

$$d\mu_\phi = \frac{80.56N_e}{f^3}df. \tag{3.22}$$

Substitution of equations 3.20 and 3.22 in equation 3.21, and simplifying gives the group refractive index as

$$\mu_g = 1 + \frac{40.28N_e}{f^2}. \tag{3.23}$$

Thus the phase and group velocities of the wave would be

$$V_\phi = \frac{c}{1 - \frac{40.28}{f^2}N_e} \tag{3.24}$$

$$V_g = \frac{c}{1 + \frac{40.28}{f^2}N_e} \tag{3.25}$$

where c is the speed of light.

Thus, conventionally, the refractive indices can be re-written as

$$n_\phi = 1 - \frac{40.28}{f^2}N_e \tag{3.26}$$

$$n_g = 1 + \frac{40.28}{f^2}N_e \tag{3.27}$$

where n_ϕ and n_g are phase and group refractive indices respectively, and will be used to derive TEC from radio signals in the following section.

3.2 TEC from ionospheric refraction

Techniques designed to infer ionospheric electron content are subject to the complex and variable refractive index of the ionosphere due to the electron density fluctuations under different conditions. Past TEC studies utilised TEC measurements derived from radio signals using the

Faraday rotation technique (e.g. Sethia et al., 1978; Davies, 1980, 1990). This technique is based on the rotation of the electric field vector of the wave resulting in polarization (Davies, 1990). This method is capable of capturing electron content that can not be determined using bottom-side sounders such as ionosondes. However, the disadvantages of the Faraday rotation method are associated with the inability to accurately know the initial direction of the electric field at the source of the signal as well as the full integer number of cycles of rotation (Davies, 1990). Details of the Faraday rotation technique's electron content dependance on the ionosphere's refractive index for both ordinary and extraordinary waves can be found in Davies (1990). The spatial and temporal characterisation of the ionosphere's behaviour for short- and long-term understanding and applications has now been shifted more to the use of electron content derived from the dispersive effect of the ionosphere on single and dual frequency signals (e.g. Komjathy, 1997; Schaer, 1999; Meggs, 2005; Yizengaw et al., 2005). Thus in this section, the derivation of TEC based on radio signal information will mainly refer to the GPS signal unless otherwise stated. As the GPS signal propagates through the ionosphere, it is bent as a result of Fermat's principle which states that: of all paths that can be taken by an electromagnetic wave (e.g. a light wave, a radio wave or a GPS signal) travelling from one point to another, it will take the path which requires the least amount of time (Hofmann-Wellenhof et al., 1992; Langley, 2000). Thus the determination of ionospheric refraction follows from the expression

$$ndx = cdt \tag{3.28}$$

where

n is the refractive index,
dx is the distance travelled by the signal,
c is the speed of light,
dt is the time taken to travel a distance dx in the ionosphere.

The time taken by the GPS signal to move from the satellite to the receiver on the ground can be determined by integration over the entire path signal. From equation 3.28,

$$t = \int_s^r \frac{n}{c} dx. \tag{3.29}$$

The measured geometric phase and group ranges are obtained by substituting the expressions for phase and group refractive indices in the expression

$$\rho = ct = c \int_s^r \frac{n}{c} dx \tag{3.30}$$

where ρ is the geometric range or distance.

Therefore

$$\rho_\phi = \int_s^r \left(1 - \frac{40.28 N_e}{f^2}\right) dx \qquad \text{phase geometric range} \tag{3.31}$$

$$\rho_g = \int_s^r \left(1 + \frac{40.28 N_e}{f^2}\right) dx \qquad \text{group geometric range.} \tag{3.32}$$

Re-writing the above equations results in

$$\rho_\phi = \int_s^r dx - \frac{40.28}{f^2} \int_s^r N_e dx \tag{3.33}$$

$$\rho_g = \int_s^r dx + \frac{40.28}{f^2} \int_s^r N_e dx \tag{3.34}$$

where

$\int_s^r dx = \rho$ (true geometric range),
$\int_s^r N_e dx =$ TEC (total electron content).

The differences between the phase/group and true geometric ranges give the ionospheric refraction

$$\rho_\phi - \rho = \Delta_\phi = -\frac{40.28}{f^2}\text{TEC} \tag{3.35}$$

$$\rho_g - \rho = \Delta_g = \frac{40.28}{f^2}\text{TEC} \tag{3.36}$$

The presence of the ionosphere therefore reduces the carrier phase measurements (phase advance) and increases the pseudo-range measurements (signal delay) by equal but opposite amounts (Hofmann-Wellenhof et al., 1992; Langley, 2000; Gao and Liu, 2002; Misra and Enge, 2006). TEC in equations 3.35 and 3.36 represent the slant TEC (along the signal path). To obtain vertical TEC, a suitable mapping function is used which takes into account the satellite zenith distance (Hofmann-Wellenhof et al., 1992), as geometrically represented in fig. 3.2. Therefore, taking into account the vertical TEC component, the phase and group geometrical ranges become

$$\Delta_\phi = -\frac{1}{\cos(z')}\frac{40.28}{f^2}\text{TEC} \tag{3.37}$$

$$\Delta_g = \frac{1}{\cos(z')}\frac{40.28}{f^2}\text{TEC}. \tag{3.38}$$

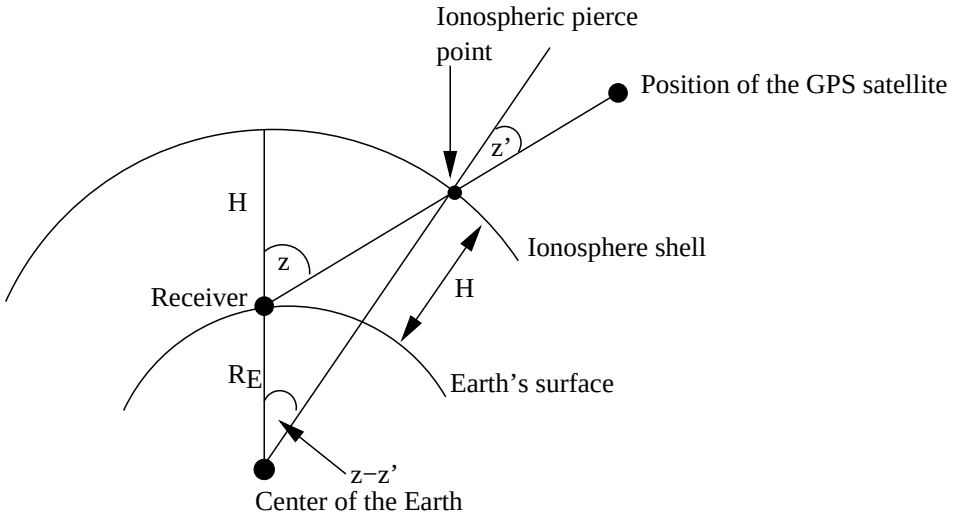


Figure 3.2: A schematic illustration of geometrical conversion of slant TEC to vertical TEC (Hofmann-Wellenhof et al., 1992).

The zenith angle (z') at the ionospheric piercing point (IPP) is computed from the expressions

$$T_s = \frac{T_v}{\cos(z')}, \quad \sin(z') = \frac{R_E}{R_E + H}\sin(z) \tag{3.39}$$

where

T_s, T_v: Slant and vertical TEC respectively,
R_E: Earth's radius (6378.134 km),
H: Single layer model height,
z, z': Zenith angles at the observation location and IPP, respectively.

In this case, the IPP is defined as the point of intersection between the GPS signal and the ionospheric shell (Hofmann-Wellenhof et al., 1992; Gao and Liu, 2002). In particular, the slant TEC is the ionospheric delay at the IPP above a certain geographic point as observed from the receiver's position, and the aim of the mapping function is to allow a geometric conversion of this slant ionospheric delay to vertical ionospheric delay at the same geographic point. The mapping function as depicted in fig. 3.2 and defined in equation 3.39 is the one that was used to convert slant TEC to vertical TEC in this study (see next section).

3.3 Estimating GNSS TEC over South Africa

3.3.1 The ASHA algorithm

This algorithm was specifically developed to estimate regional TEC values over South Africa (Opperman et al., 2007). It was adapted from the Conventional Spherical Harmonic Analysis

(CSHA) procedure developed by Schaer (1999). In his work, Schaer (1999) pointed out that the global model could be adapted for regional applications. During the ASHA development process, a series of validations was carried out using other independent data sources such as ionosonde, IRI model, Multi-Instrumental Data Analysis System (MIDAS), as well as TEC derived from Global Ionospheric Maps (GIM), all of which agreed favourably with regard to diurnal variability. Over particular locations such as the Louisvale ionosonde station, some differences in the performances of different methods were noted, such as the existance of peaks in ASHA and MIDAS TEC that were not evident in GIM. These differences are partially attributed to data resolutions used for these methods. In the validation study, ASHA, MIDAS and GIM used time resolutions of 1 minute, 15-30 minutes and 2 hours respectively (Opperman et al., 2007; Opperman, 2008). Like the global CSHA, the ASHA algorithm does TEC estimation and modelling using the spherical harmonic expansion (Schaer, 1999; Opperman et al., 2007)

$$TEC(\lambda,\phi) = \sum_{n=0}^{N}\sum_{m=0}^{n} \overline{P}_{nm}\left\{\cos(\phi)\right\}\left[a_{nm}\sin(m\lambda) + b_{nm}\cos(m\lambda)\right] \tag{3.40}$$

where

λ: IPP sun-fixed longitude $\in [0^\circ, 360^\circ]$,
ϕ: IPP co-latitude $\in [0^\circ, 180^\circ]$,
$\overline{P}_{nm}$: Normalised associated Legendre functions,
a_{nm}, b_{nm}: The unknown spherical harmonic coefficients to be estimated from a weighted least squares solution,
n and m are the degree and order of the spherical expansion respectively.

Both the original CSHA and the ASHA algorithms use the IPP sun-fixed longitude which takes into account the IPPs time and longitudinal parameters in a single angular observation over a span of 360° for the entire 24 hr revolution (Opperman et al., 2007; Opperman, 2008) and is computed from the geographical longitude λ_g as follows;

$$\lambda = (180^\circ - \Omega_e.t) - \lambda_g \tag{3.41}$$

where λ is the sun fixed longitude, Ω_e is the Earth rotation rate and t is the elapsed time (in universal time) since midnight. The λ is, therefore, the longitude expressed relative to the sun's mean geographic longitude.
While no modifications are done for the co-latitude in the CSHA (since the n wavelengths of the Legendre polynomial are effectively defined over 180° that almost covers the IPPs within the global international GNSS service (IGS) network), in the ASHA algorithm (which is regionally based) scaling is required, since the latitudinal span from the South African GPS network is about 30°. This is a narrow latitude band in the sun-fixed longitude system (Opperman et al.,

2007). This is done by translating and scaling the narrow co-latitude band of the IPPs to a hemisphere with the minimum co-latitude and the latitudinal span of the GPS observations (De Santis et al., 1991; Opperman et al., 2007)

$$\phi' = \frac{90^\circ}{\theta}(\phi - \phi_o) \tag{3.42}$$

where
ϕ', ϕ_o and θ are the scaled co-latitude, minimum co-latitude and latitude span respectively.
For regional modelling, this scaled co-latitude effectively replaces ϕ in equation 3.40. The ASHA algorithm estimates slant TEC from pseudo-range GPS measurements and uses it to eliminate the ambiguity offset in the calculated carrier phase TEC by a process known as "levelling". This combination of measurements eliminates a large percentage of ionospheric error and gives TEC to a better accuracy. To minimise the multipath errors, an elevation cutoff angle of 15 degrees is used. The slant TEC estimated from GPS observations is converted to vertical TEC by assuming the ionosphere to be a single layer height of 350 km.

3.3.1.1 The simplified algorithm

As previously mentioned, the general ionospheric delay is directly proportional to TEC and inversely proportional to the square of the frequency as follows:

$$\Delta\tau = \left(\frac{40.28}{cf^2}\right) \times TEC = \alpha.T_s, \qquad \alpha = \frac{40.28}{cf^2}, \tag{3.43}$$

where T_s is the slant TEC along the ray path of the signal through the IPP. The vertical ionospheric delay is given by

$$\Delta\tau = \left(\frac{40.28}{cf^2}\right)\left(\frac{T_v}{cos(z')}\right) + (b^r - b^s) = \alpha\beta.T_v + (b^r - b^s) \tag{3.44}$$

where
T_v is the vertical TEC, $\beta = 1/cos(z')$, z' is the zenith angle at the IPP, b^r and b^s are the receiver and satellite interfrequency differential biases respectively.
Writing vertical TEC in terms of spherical harmonic expansion,

$$T_v(\lambda, \phi) = \sum_{n=0}^{N}\sum_{m=0}^{n} \overline{P}_{nm}\left\{\cos(\phi)\right\}\left[a_{nm}\sin(m\lambda) + b_{nm}\cos(m\lambda)\right].$$

Vertical TEC can be expressed in terms of its components as

$$T_v(\lambda,\phi) = \sum_{n=0}^{N}\sum_{m=0}^{n} \overline{P}_{nm}\left\{\cos(\phi)\right\}.a_{nm}\sin(m\lambda) + \sum_{n=0}^{N}\sum_{m=0}^{n} \overline{P}_{nm}\left\{\cos(\phi)\right\}.b_{nm}\cos(m\lambda). \quad (3.45)$$

By substitution of equation (3.45) in equation (3.44), the ionospheric delay becomes

$$\Delta\tau = \alpha\beta\left(\sum_{n=0}^{N}\sum_{m=0}^{n} XY + \sum_{n=0}^{N}\sum_{m=0}^{n} XZ\right) + (b^r - b^s) \quad (3.46)$$

where

$$X = \overline{P}_{nm}\left\{\cos(\phi)\right\}, \quad Y = a_{nm}\sin(m\lambda) \;\; and \;\; Z = b_{nm}\cos(m\lambda).$$

The Jacobian method is applied where partial derivatives of ionospheric delay with respect to the coefficients a_{nm} and b_{nm} and the biases b^r and b^s are taken, thereby estimating them:

$$J(a_{nm},\ b_{nm},\ b^r,\ b^s) = \left(\frac{\partial\Delta\tau}{\partial a_{nm}} \quad \frac{\partial\Delta\tau}{\partial b_{nm}} \quad \frac{\partial\Delta\tau}{\partial b^r} \quad \frac{\partial\Delta\tau}{\partial b^s}\right). \quad (3.47)$$

In the ASHA algorithm, the differential clock biases (DCBs) from both satellite transmitters and GPS receivers are estimated (and eliminated) along with the spherical harmonic coefficients using the least squares method. The estimated coefficients are then used in the spherical harmonic expression to estimate the ionospheric delay. A full solution of this is discussed in Opperman et al. (2007) and Opperman (2008).

3.3.2 The UNB-IMT model

The UNB-IMT is an algorithm that was developed by Komjathy (1997) to estimate global and regional TEC values with the purpose of providing corrections for surveillance, navigation as well as communication applications. Above each GPS receiver in the network considered, the UNB algorithm employs a spatial linear approximation of stochastic parameters in a Kalman filter optimisation to describe TEC dependance on local time and geomagnetic latitude (Komjathy, 1997). It uses phase levelling as a combination of both carrier-phase and pseudo-range measurements in TEC estimations. The UNB algorithm also uses the thin-shell mapping function based on the IRI derived variable shell height to transform slant TEC to vertical TEC and estimates this ionospheric delay according to the following expression

$$I_{r_j}^{s_i}(t_k) = M(e_{r_j}^{s_i}).\left\{a_{0,r_j}(t_k) + a_{1,r_j}(t_k)\delta\lambda_{r_j}^{s_i} + a_{2,r_j}(t_k)\delta\phi_{r_j}^{s_i}\right\} + b_{r_j} + b^{s_i} \quad (3.48)$$

where

$j = 1, 2, ..., n$ represents the number of receivers in a given network,

$I_{r_j}^{s_i}(t_k)$ is the phase-levelled $L1-L2$ TEC measured by receiver r_j observing satellite s_i at time t_k,
$M(e_{r_j}^{s_i})$ is the thin shell mapping function that transforms slant TEC to vertical TEC and is defined as

$$M(e_{r_j}^{s_i}) = \left(1 - \frac{r_E \cos^2 e_{r_j}^{s_i}}{r_E + h}\right)^{-\frac{1}{2}}$$

with r_E being the mean radius of the Earth and h the mean value for the assumed height of the thin spherical shell (400 km) (Komjathy, 1997),
$a_{0,r_j}, a_{1,r_j}, a_{2,r_j}$ are the parameters for spatial linear approximation of TEC that are estimated per GPS station r_j,
$e_{r_j}^{s_i}$ is the elevation angle of satellite s_i observed by receiver r_j,
$\delta\lambda = \lambda_{r_j}^{s_i} - \lambda_0$ is the longitude difference between the IPP and that of the mean sun,
$\delta\phi = \phi_{r_j}^{s_i} - \phi_{r_j}$ is the geomagnetic latitude difference between the IPP and that of the receiver,
b_{r_j} and b^{s_i} are the receiver and satellite differential delays respectively.
Similar to other approaches (e.g. Mannucci et al., 1995; Schaer, 1999; Opperman et al., 2007), the UNB technique uses the solar geomagnetic reference frame to compute TEC since the ionospheric properties are based mainly on the interactions between the solar radiation and the Earth's geomagnetic field (e.g. Komjathy, 1997). In TEC mapping, this method gives hourly TEC maps using a $1^\circ \times 1^\circ$ grid (for regional mapping) and a $5^\circ \times 5^\circ$ grid (for global mapping). This technique was adopted to study solar cycle effects of GNSS-derived TEC over Southern Africa using both IGS receiver stations and the network of receivers operated by South Africa's Chief Directorate Surveys and Mappings (CDSM) near both solar maximum and minimum for solar cycle 23 (Moeketsi et al., 2007; Moeketsi, 2008).

3.4 Summary

This chapter gave a brief introduction to GPS measurements. The discussion of ionospheric effects on GPS measurements along with the derivation of group delay and phase advance introduced into GPS signals was presented. The ASHA algorithm, which was used to derive TEC values over South Africa was described as well as the UNB-IMT algorithm. In this study TEC derived values from the UNB-IMT were compared to the developed simulation model's TEC predictions. A comparison of TEC values computed using the ASHA, UNB-IMT algorithms and their corresponding predicted counterparts (IRI-2007 and developed NN models) will be performed and details can be found in chapter 6. This was done towards satisfying one of the objectives of this research which was to show that a TEC simulation model could be used to complement these algorithms (which are mainly satellite data dependent) in generating complete datasets over Southern Africa.

Chapter 4

A feasibility study

4.1 Introduction

This chapter describes a feasibility study carried out to investigate the possibility of modelling TEC values over a particular GPS receiver station using the NN and regression models developed with GPS data obtained from a different receiver station. It also includes the comparison of GPS TEC and TEC values generated by the International Reference Ionosphere (IRI) for South Africa. The main objective was to establish the suitability of NNs and regression analysis in modelling TEC values for locations where data was not included in the models' development. In addition, this work provides background material for the quest towards the development of a TEC prediction model for Southern Africa. In short, this chapter presents results obtained by using separate NN and regression models developed over GPS stations located at Sutherland (32.38°S, 20.81°E) and Springbok (29.67°S, 17.88°E), in order to model TEC variability over Cape Town (33.95°S, 18.47°E) and Upington (28.41°S, 21.26°E) respectively. Some results from this chapter also appear in the publication by Habarulema et al. (2009a).

4.2 TEC variability

As shown in chapter 2 the variability of the ionosphere is influenced by the sun. From this theory it is possible to deduce parameters that should be considered when predicting TEC variability. The parameters covered in this section are those known to influence TEC variability, taking into account the intention to develop a local TEC model for a mid-latitude region.

4.2.1 Solar activity

The sun is the major source of energy that leads to photoionisation, thereby contributing greatly to the electron density within the ionosphere. It exhibits a cycle of $\sim$ 11 years displaying both

high and low solar activity periods. Ultimately, TEC varies with the activity of sun, its highest values occuring during the maximum solar activity periods and lowest during solar minimum. The sunspot number is one of the most commonly used indicators of solar activity and hence is a parameter used for input when modelling with NN or regression analysis.

4.2.2 Magnetic activity

The ionosphere's complex behaviour is different depending on the latitudinal location over which its observations are made. The ionospheric variability is influenced by the Earth's magnetic field at low, mid and high latitudes (e.g McNamara, 1991). A magnetic activity parameter (denoted by K) is the measure of the geomagnetic field variations over any location on the globe as a result of currents flowing in the ionosphere and the magnetosphere. Ultraviolet radiation from the sun is the source of thermal convection at ionospheric heights and this leads to the movement of ions and electrons accross the geomagnetic field. The generated ionospheric current gives rise to a magnetic field in the ionosphere's surroundings the variations of which are later observed as geomagnetic field fluctuations on the Earth's surface (Ondoh and Marubashi, 2000). To take into account the effect of TEC variations with Earth's geomagnetic field in NN modelling, the magnetic A index values derived from the K index recorded at the Hermanus Magnetic Observatory in South Africa was included as an input.

4.2.3 Diurnal variation

The ionosphere varies throughout the day with the maximum electron density expected at about local midday. This variation directly translates into fluctuation of ionospheric parameters, including TEC. TEC reaches a maximum between $\sim$ 12h00 to 14h00 South African Standard Time (SAST=UT+2) during the day, where UT is the universal time. Due to the disappearance of some ionospheric layers from the bottomside during the night (D, E and F1), the ionospheric plasma reduces and hence low TEC values are expected during the nighttime. This is generally due to the absence of photoionisation. The UT hour values were used to represent the diurnal variation of TEC as an input to the NN and regression models.

4.2.4 Seasonal variation

The solar zenith angle (in this case, the angle between the point directly over the GPS station (zenith) and the direction of the sun) varies both diurnally and seasonally (McNamara, 1991). Seasonal variation simply refers to the day to day variability throughout the whole year. Due to decreased photoionisation levels, the electron density in winter has low values compared to summer, and thus TEC is expected to exhibit a similar behaviour. Therefore for statistical models where examples are required for the NN to learn the variational patterns of TEC over

a year and for analytical regression models to exhibit the seasonal character, the day number is sufficient and was hence included as an input.

4.2.5 Location (geographic latitudes and longitudes)

The variation of the ionosphere with position on the Earth is due to solar zenith angle variations (McNamara, 1991) which later determine the ionisation levels. The ionospheric electron density varies considerably with latitude and therefore for spatial characterisation and generalisation of TEC change patterns, the latitude and longitudes were included in TEC modelling. It will however be shown that, since this feasibility study uses single station data, the geographical aspect was not included at this stage.

4.3 Modelling techniques

In this chapter two techniques, namely regression analysis and NNs, were applied to the same dataset to model TEC variability as a function of similar parameters. The concept of NNs was explained in chapter 2 and therefore only regression analysis will be briefly described here. In regression analysis the parameter to be modelled is known as a dependent variable and is estimated as a function of independent variables. From a single station modelling point of view, taking TEC as the target value, the resulting regression function comprises representation of seasonal and diurnal variations, and solar and magnetic activity. Two regression analysis options (simple and multiple linear regression techniques) were explored for the TEC estimations.

4.3.1 Simple regression analysis

In simple regression analysis, the dependent variable is modelled as a linear combination of independent variables. Generally, the modelled parameter is a function of independent variables and unknown coefficients, i.e.

$$\mathrm{y} = f(\mathbf{X}, \beta)$$

where $\mathbf{X}$ represents the independent variables and in this case, parameters that influence TEC variability and β are the unknown parameters or coefficients of the regression function which are determined by the minimization of the squares of the deviations of the data from the model using the equation

$$\mathbf{X_m}\beta = \mathbf{Y} \tag{4.1}$$

where $\mathbf{X_m}$ is the matrix comprising of independent variables and $\mathbf{Y}$ is a column matrix of the dependent variable. The resulting equation for modelling a particular parameter is of the form

$$y = \beta_0 + \beta_1 x_1 + ... + \beta_n x_n \tag{4.2}$$

where $x_1, x_2, ...x_n$ are the parameters that influence TEC variability.

4.3.2 Multiple regression analysis

A regression method that models the dependent variable as a product of independent variables raised to unknown powers was employed, i.e.

$$y = C x_1^{\beta_1} \times x_2^{\beta_2} \times ... \times x_n^{\beta_n} \tag{4.3}$$

The foundation of the above expression is linear regression analysis by means of the linear combination of logarithms of the considered parameters (e.g. Verő, 1980; Heilig et al., 2010) as,

$$\log(y) = d + \beta_1 \log x_1 + \beta_2 \log x_2 + ... + \beta_n \log x_n \tag{4.4}$$

where $d = \log C$. The underlying idea is that given the independent variables, $x_1, x_2, ..., x_n$, the dependent variable can be estimated from equations 4.2 and 4.3.

4.4 Data processing for TEC modelling

A knowledge of ionospheric variability indicates that multiple inputs should be used for the NN to learn and store information about TEC variability which can later be used for generalisation. The same inputs or parameters were also used in regression models to generate analytical formulae for TEC estimations. GPS TEC values were calculated from the dual frequency GPS receiver network using the ASHA algorithm at intervals of 60 s. For the purpose of this preliminary work, low resolution data (hourly values of TEC) were extracted from the data set generated by the ASHA algorithm and used in NN and regression modelling. The dataset used to train the NN and derive regression functions consisted of the TEC values at each hour for the period 2001-2004. A reduced dataset was deliberately used for the feasibility study in order to reduce the training time of the NN. The inputs for the NN and regression models were chosen from the parameters known to influence TEC, such as solar activity, magnetic activity, seasonal variation, diurnal variation and geographical location of the GPS receivers. The first four parameters were considered since the feasibility study only involved single station studies. Fig. 4.1 shows the generalised picture of the procedure that was followed for TEC modelling in this chapter. The representations of solar and magnetic activity variations for TEC

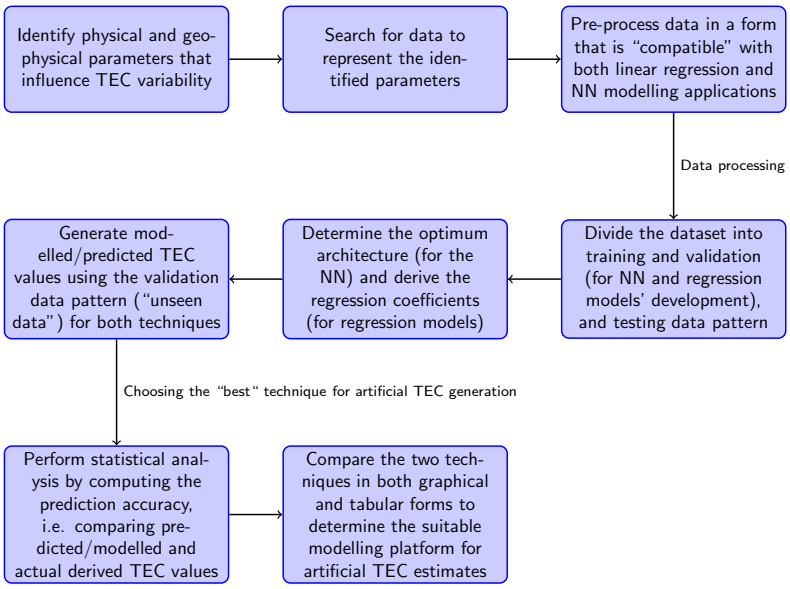


Figure 4.1: A simplified flow diagram of the TEC modelling procedure in this chapter.

modelling were determined by training the NN model with the 2000-2004 dataset at 10h00 UT over the Square Kilometre Array (SKA) Hub location (30.71°S, 21.39°E), South Africa, with different combinations of solar and magnetic activity variables. The training dataset consisted of a combination of seasonal variation (represented by day number), and solar and magnetic activities for which suitable representations had to be determined. Training was done using data for R0, R1, R2, ... , R12 with each of A4, A8, A16, ... , A128. R0 and R1 are defined as the daily sunspot number and 1-month running mean of the daily sunspot number respectively. R2, R3, ..., R12 take on similar appropriate definitions. The A-index was obtained from the 3-hourly values of the K-index. Since for each recorded K-index value, there is a corresponding a_k value, A4 is defined as the value corresponding to the mean (average) of the previous 4 a_k indices and A8, ..., A128 follow the same definition pattern (Habarulema et al., 2007). These magnetic A index values were derived from the K-index recorded at the Hermanus Magnetic Observatory, South Africa to ensure that the response of TEC is clearly linked to the local magnetic conditions. Generally, statistical methods are used to determine the NN optimum solution in empirical modelling. In this thesis, the root mean square error (RMSE) method was used to identify the model that provides suitable TEC estimates, and is defined as:

$$\text{RMSE} = \left(\frac{1}{N} \sum_{i=1}^{N} (vT_m - vT_p)^2 \right)^{\frac{1}{2}} \tag{4.5}$$

where $i = 1, 2, ..., N$, N is the total number of observations for a given dataset, and vT_m and vT_p are the measured and predicted vertical TEC values respectively. This method (RMSE) has been widely used as a means to determine optimum parameters for ionospheric peak electron

GPS station	geographic latitude (°S)	geographic longitude (°E)
Sutherland	32.38	20.81
Cape Town	33.95	18.47
Springbok	29.67	17.88
Upington	28.41	21.26
SKA Hub	30.71	21.39

Table 4.1: GPS stations and their geographical positions used in this preliminary study

density predictions (Williscroft and Poole, 1996; McKinnell, 2002; Oyeyemi et al., 2006), solar and geomagnetic activity data predictions (Macpherson et al., 1995), solar cycle predictions using NNs (Conway et al., 1998), modelling of solar wind-magnetosphere interaction (Wu and Lundstedt, 1997) and many other statistically based predictions. Fig. 4.2 illustrates the RMSE values obtained from each of the NNs trained during the search for suitable representations for solar and magnetic activities. From this figure, it is clear that the NN with R4 and A8 provide the optimum solution and was hence adopted for TEC predictions and modelling over South Africa. A detailed discussion about the determination of appropriate input parameters to represent solar and magnetic activity over South Africa suitable for TEC modelling is found in Habarulema et al. (2007). The seasonal and diurnal variations, as represented by day number (D_n) and hour (H_r), were each split into two cyclical components to allow a numerical continuous trend of the data (Williscroft and Poole, 1996; Poole and McKinnell, 2000) as follows:

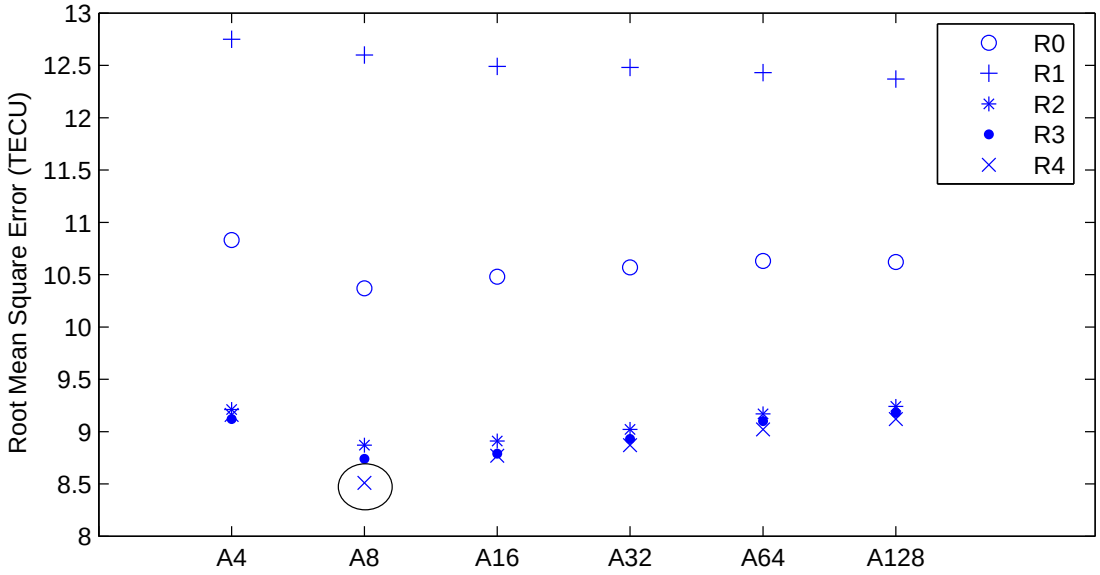
$$D_{ns} = \sin\left(\frac{2\pi \times D_n}{365.25}\right) \quad D_{nc} = \cos\left(\frac{2\pi \times D_n}{365.25}\right) \tag{4.6}$$

$$H_{rs} = \sin\left(\frac{2\pi \times H_r}{24}\right) \quad H_{rc} = \cos\left(\frac{2\pi \times H_r}{24}\right) \tag{4.7}$$

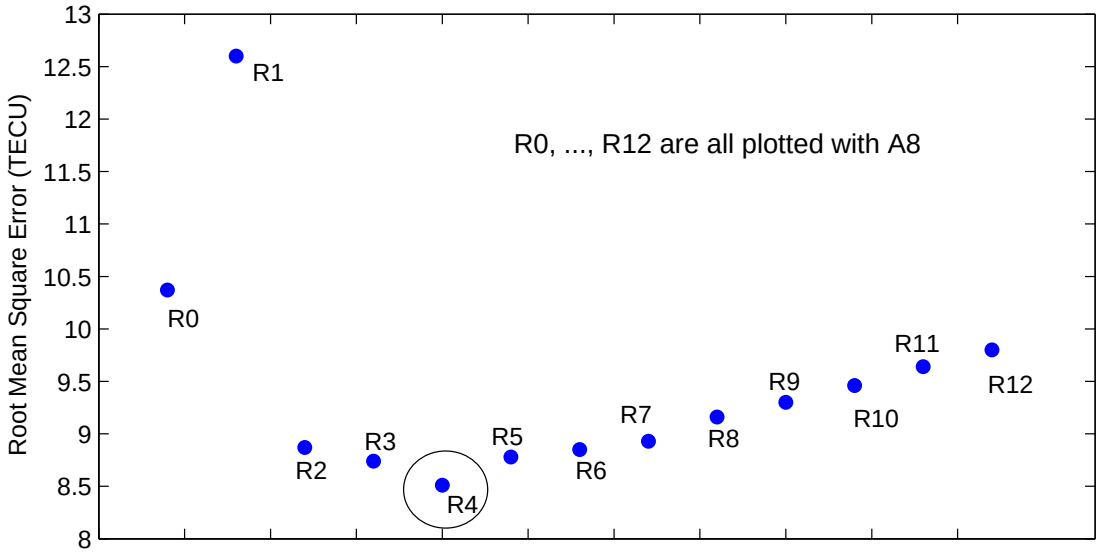
where

D_{ns}, D_{nc}, H_{rs} and H_{rc} are the sine and cosine components of D_n and H_r respectively. The factor 0.25 in the above equations (equation 4.6) take into account the fact that there are some data for leap years. Table 4.1 shows the GPS stations used with their geographical positions. Fig. 4.3 shows the map of South Africa depicting the approximate geographical locations of the GPS receivers used in this preliminary study.

For NN and regression modelling the target values are the vertical TEC derived from the slant TEC (TEC along the line of sight) at the IPP, assuming that the ionosphere is a single layer (usually referred to as the Single Layer Ionospheric Model, SLIM). The latter is described in chapter 3. Ideally, for NN modelling, the dataset is divided into the training, validation/verification and testing sets.



(a) Computed Root Mean Square Errors (RMSEs) between GPS TEC and NN TEC (TECU) values for the determination of the appropriate representation for the magnetic activity. The previous eight 3-hourly magnetic A index values (A8) gave the least RMSE and was hence chosen as the optimum.



(b) A plot of RMSEs between GPS TEC and NN TEC (TECU) values for R0, R1,..., R12 with the running mean of the previous eight 3-hourly magnetic A index values (A8). R0 is the daily sunspot number, R1 is 1-month running mean of daily sunspot number and R2, R3, ..., R12 are defined in a similar way.

Figure 4.2: RMSE values for the NNs trained with different combinations of input parameters. Circled points represent the combination R4A8 that gave the optimum solution (Habarulema et al., 2007).

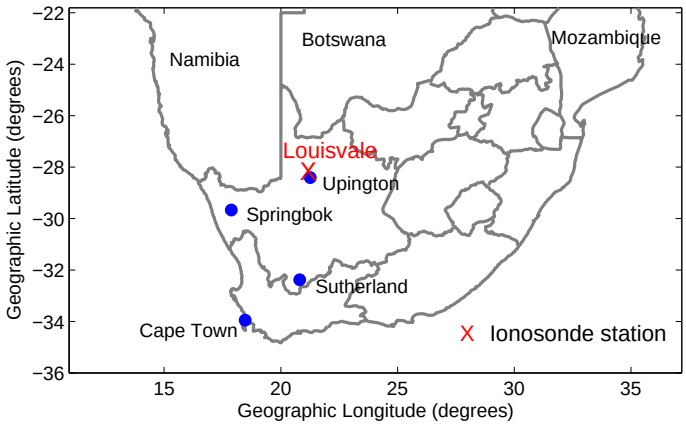


Figure 4.3: A South African map showing the four dual frequency GPS receiver stations and the location of the ionosonde used in this study.

- The training dataset is the one used for training the NN by adjusting the weights between input, hidden and output layers.
- The validation/verification dataset determines the performance of the NN on patterns that are not trained during learning.
- The testing set is used to assess the final overall performance of the NN.

The dataset comprising both training and verification datasets forms mainly the basis of the NN model. These datasets were used for the development of the NN and regression models for fair comparisons, while the developed models were verified using a similar testing dataset.

4.5 The NN implementation

As already defined in chapter 2, a NN is an interconnected group of artificial neurons that can be applied to learn data patterns from both known input and output parameters using computational modelling (Fausett, 1994). A NN consists of input, hidden and output layers. In each layer there are nodes or units that calculate numerical values which are dependant on the number of units present in the entire network connection (Conway et al., 1998). The main advantage of NNs is that given enough historic data, they are capable of learning and generalising dependencies between unknowns which exhibit non-linear behaviour when identical but not necessarily the same data is introduced (Fausett, 1994; Bishop, 1995). NNs have different forms and a wide range of applications. These are discussed in several sources (Haykin, 1994; Fausett, 1994; Bishop, 1995). For this feasibility study, a single hidden layered feed forward network (SFFNN) with a back propagation algorithm was used. This SFFNN has one input layer, one hidden layer and one output layer. The choice of this network is based on the

work done by Habarulema et al. (2007) which describes the first attempt at predicting TEC over South Africa using NNs and established that they were indeed suitable for predicting TEC derived from the South African GPS receiver network. It has also been shown that including more than one hidden layer does not lead to much difference in the accuracy of results (Haykin, 1994), although more hidden layers may make training easier in some cases (Fausett, 1994), depending on the amount of data under consideration. The input layer consists of six nodes or neurons corresponding to six inputs (in this work). Using the procedure of adding a neuron to the hidden layer, training the NN, testing it and finally comparing the predicted TEC with the derived GPS TEC, the number of hidden neurons that gave the optimum solution was determined to be nine. Thus the NN configuration used in this study is 6:9:1, where one is the neuron within the output layer.

4.6 Regression and NN models compared

Regression and NN models were developed using hourly values of TEC from a receiver station located at Sutherland for the period 2001-2004. The previous section described the NN implementation and the final configuration for the NN model. By following the procedure in section 4.3, the simple linear regression (SLR) model for modelling TEC variability with an hourly interval during 2001-2004 was found to have the function

$$y = 3.6068 + 0.4372D_{nc} + 6.6253D_{ns} + 1.199H_{rs} - 18.3787H_{rc} + 0.2428R4 - 0.0994A8. \quad (4.8)$$

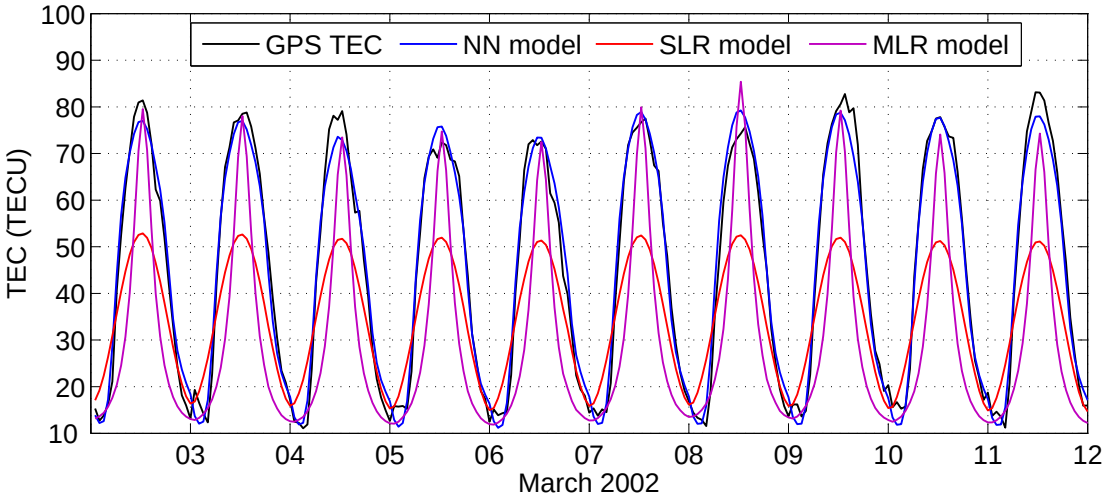
The multiple linear regression (MLR) model for estimating TEC over the same location using the same dataset is

$$y \sim 0.834(D_{ns}{+}a)^{-0.0098}{\times}(D_{nc}{+}a)^{0.4081}{\times}(H_{rs}{+}a)^{-0.0068}{\times}(H_{rc}{+}a)^{-0.7569}{\times}(R4)^{0.6948}{\times}(A8{+}b)^{-0.0831} \quad (4.9)$$

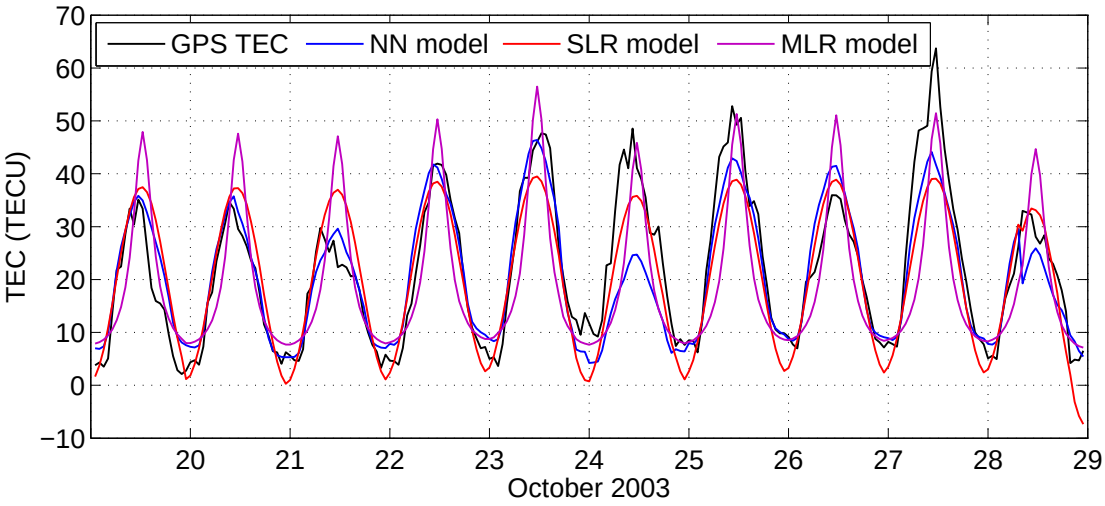
where y is TEC, $a = 1.2$ and $b = 0.1$. The latter are constants added to the sine and cosine components of day number and hour; and A8 respectively to allow their respective logarithms to be taken.

To analyse the performances of both regression and NN models, each model was verified using the same hourly TEC dataset over Cape Town for the duration 2002-2004 and table 4.2 shows the correlation coefficients computed for the three models.

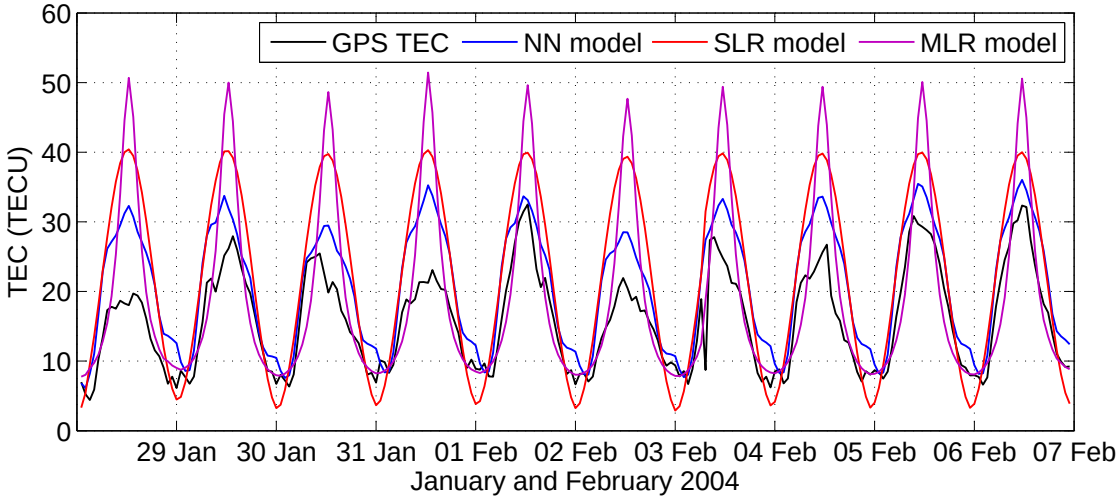
Fig. 4.4 shows the graphical comparison of modelled hourly TEC values using the NN, SLR and MLR models and GPS TEC over Cape Town for 03-12 March 2002, 20-29 October 2003 and 29 January to 07 February 2004. Table 4.2 and visual inspection of fig. 4.4 indicate that the NN model provides better estimates of TEC than the regression models. There are few cases where the SLR model tends to provide good estimations, but these may not be reliable



(a) Comparison of NN and regression models for 03-12 March 2002.



(b) Comparison of NN and regression models for 20-29 October 2003.



(c) Comparison of NN and regression models for 29 January to 07 February 2004.

Figure 4.4: Regression models and the NN model compared with GPS TEC over CPTN for selected periods in 2002, 2003 and 2004.

Table 4.2: Computed correlation coefficients between hourly measured TEC and modelled TEC (by the NN, SLR and MLR models) for 2002, 2003 and 2004 over Cape Town (33.95°S, 18.47°E).

Year	correlation coefficients		
	NN model	SLR model	MLR model
2002	0.9619	0.8655	0.8063
2003	0.9402	0.8665	0.8054
2004	0.8663	0.8007	0.7418

since this model also shows negative values for TEC (fig. 4.4(b)), which is not physically feasible, as TEC under normal conditions can not be negative. To investigate their forecasting capabilities, the developed models were used to reproduce TEC patterns for 01-09 January 2005 over Sutherland (a station used in the models' development) and compared to measured data, as shown in fig. 4.5.

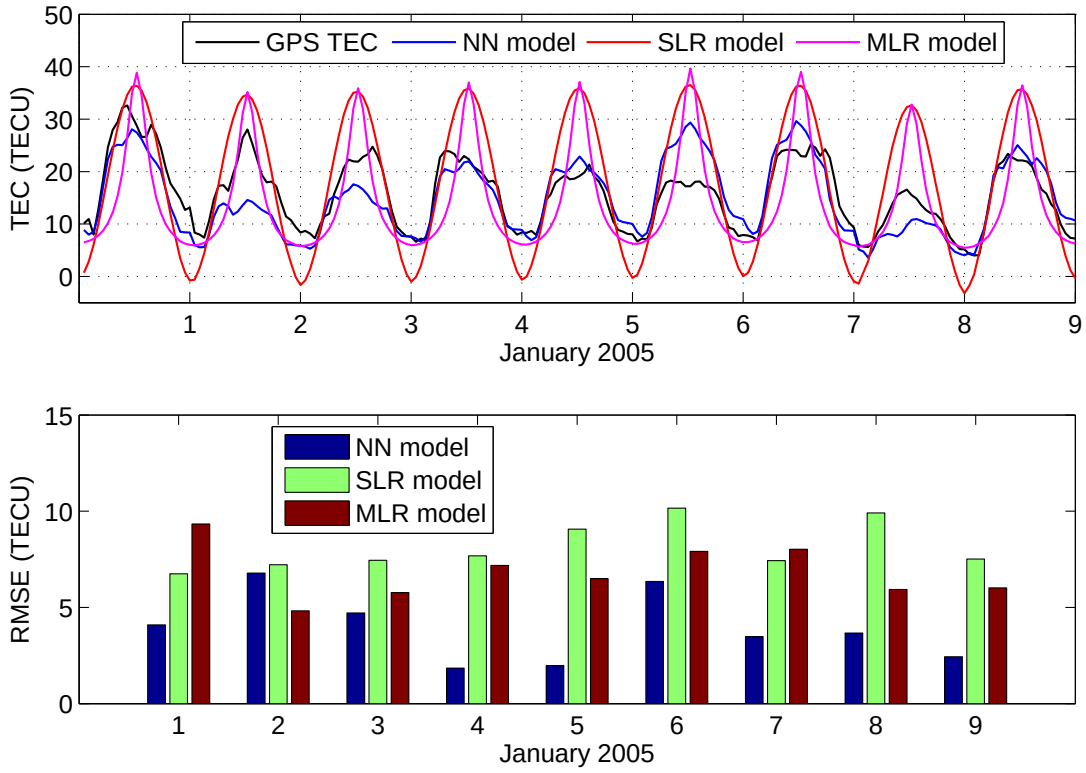


Figure 4.5: Hourly forecasted TEC values and RMSE values for 01-09 January 2005 over Sutherland.

Note the high TEC values ($\sim$ 10 TECU) during the night (in fig. 4.5) which indicates a large plasmaspheric electron content. Comparing this to values indicated in fig. 4.4, it is unrealistic

to have high TEC values at night during periods of low solar activity than high solar activity. This could be related to the poor receiver/antenna performance or the incorrect application of receiver/satellite biases as the receiver over Sutherland experienced problems prior to the period analysed. Fig. 4.6 shows mean monthly GPS TEC variations (with error bars inserted to represent standard deviation) over Cape Town (2002-2004) and Sutherland (2005), along with values computed from NN and regression models. The decrease in mean monthly GPS TEC values with a decline in solar activity is evident and well-captured by the two modelling techniques. Computed RMSE values between GPS TEC and estimates provided by the NN, SLR and MLR models (as shown in fig. 4.5) confirm that the NN model provides more accurate estimates than the regression models, except for 2 January 2005 where the NN model strongly underpredicted TEC. NNs are more linear, a probable reason why they are better able to estimate the non-linearity of TEC than the regression models developed. Due to these results, the remainder of this thesis' discussion concentrates on TEC modelling using NNs.

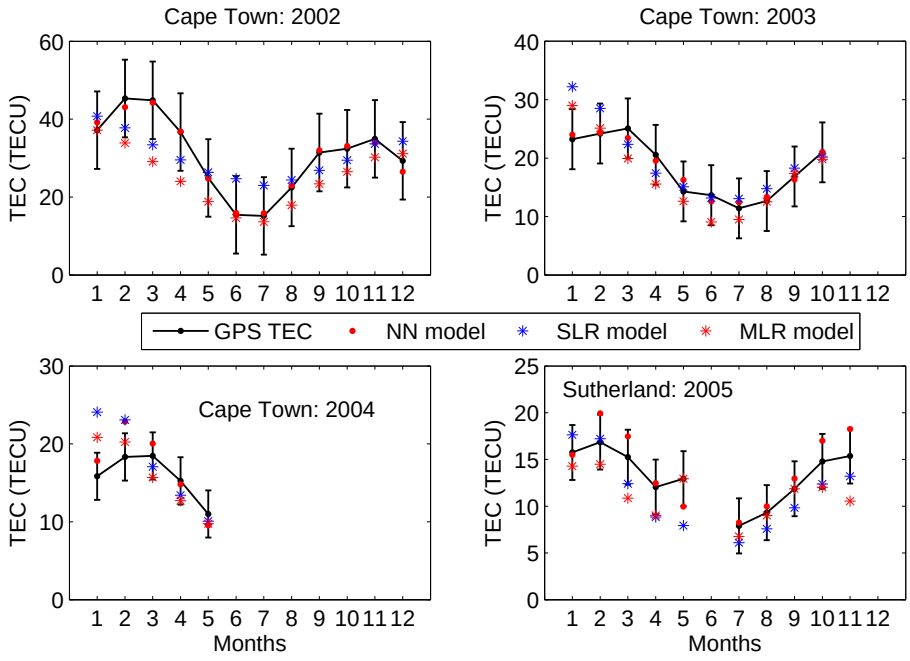


Figure 4.6: Mean monthly GPS TEC variations (2002-2005) with corresponding values computed from data generated by both modelling techniques.

4.7 NN models' results and discussion

As mentioned in section 4.5, different networks were trained in search for the architecture that provides the optimum solution. For each NN trained the RMSE values between the GPS TEC and the predicted TEC were computed for selected datasets. Based on this method, the architecture corresponding to the optimum solution was used to generate predicted TEC values

and compared with TEC values generated by the IRI. It should be noted that the IRI produces monthly median values for TEC and the NN provides the best possible average predicted TEC for the given geophysical conditions. The results were compared with the most commonly used empirical model namely the IRI-2001 predictions. In cases where the GPS receiver is near or co-located with the ionosonde, GPS TEC data were compared with ionosonde TEC (ITEC) for validation purposes. In South Africa the ionosonde data provides the best available means of validating GPS data and hence GPS-based predictions/modelled data.

4.7.1 Sutherland and Cape Town

Hourly values of TEC were extracted for a period of four years (2001, 2002, 2003 and 2004) from the dataset and used in the NN training and validation processes. The total number of inputs were 6 ($D_{ns}, D_{nc}, H_{rs}, H_{rc}$, R4 and A8) and the architecture of the network was single layered with nine hidden nodes in the hidden layer. Each input parameter was assigned an input node

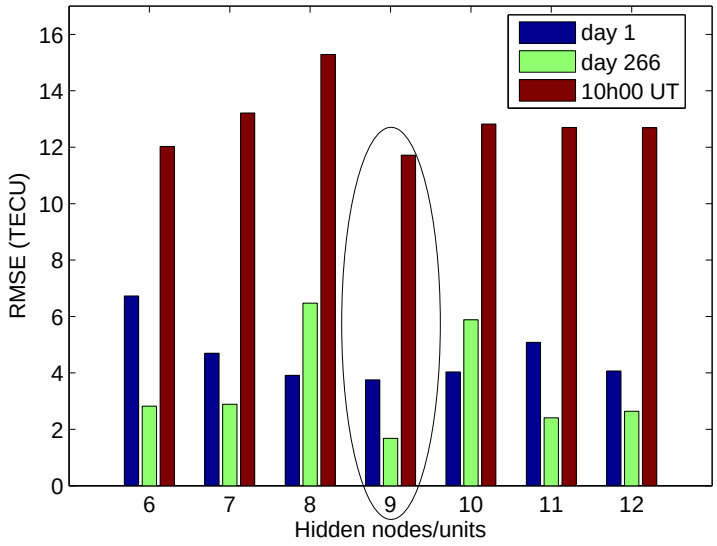


Figure 4.7: An example showing RMSE values for varying hidden nodes in search for the optimum solution over Sutherland for the NN model verified on TEC data from 2000.

and the number of hidden nodes varied followed by training the corresponding network and finally computing the RMSE values between GPS TEC and modelled TEC. An example of the network that provided the optimum solution is shown for days 1 and 266 and hourly local noon values over Sutherland for the year 2000 in fig. 4.7. The low RMSE values indicate the average close approximation of predicted values to the GPS TEC values. The constructed hourly NN model for Sutherland was used to predict GPS TEC values for the GPS receiver station located at Cape Town (33.95°S, 18.47°E) during equinoxes and solstices as shown in fig. 4.8.

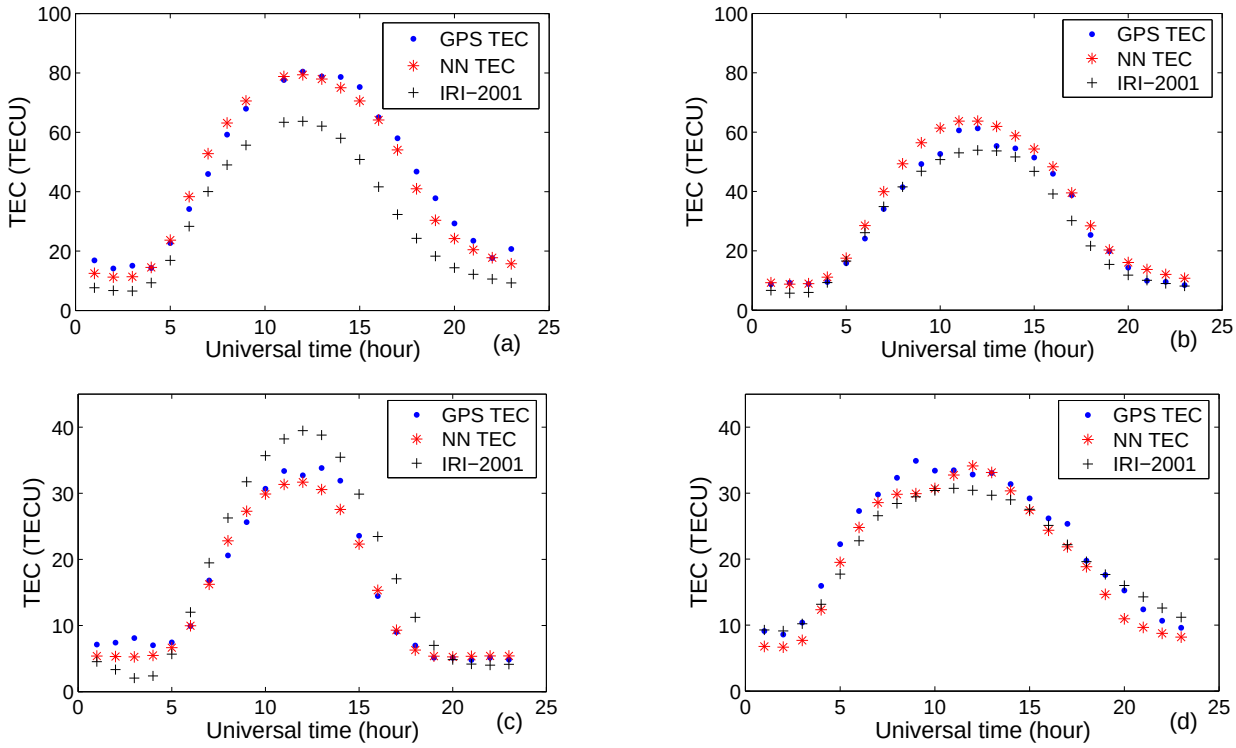


Figure 4.8: Comparison of GPS TEC with TEC predictions from the IRI-2001 and NN model over Cape Town (33.95°S, 18.47°E) during equinoxes: (a) March 20, (b) September 23 and solstices: (c) June 22, (d) December 21, all in 2002.

Fig. 4.8 shows the comparison of hourly GPS TEC values (from the ASHA algorithm) with the corresponding IRI-2001 and NN TEC predictions for days representing equinoxes and solstices over a mid-latitude GPS station located at Cape Town (33.95°S, 18.47°E). Note that IRI-2001 underpredicts GPS TEC on day 79 and overpredicts GPS TEC on day 173 (between $\sim$ 08h00 UT - 17h00 UT) in 2002 as shown in figs. 4.8(a) and (c). There is a maximum difference of $\sim$ 17 TECU between GPS TEC and IRI TEC at 12h00 UT as seen in fig. 4.8(a). This may partly be due a contribution by the plasmasphere to GPS TEC at this particular time. The remaining differences may be due to other sources such as errors introduced into the GPS TEC value by assumption that the ionosphere is a single layer of fixed height and the assumptions made about the topside profile in the IRI model. GPS TEC values compare favourably with IRI TEC values in figs. 4.8(b) and (d), with maximum differences of $\sim$ 8 TECU and 5 TECU observed at 11h00 UT and 09h00 UT respectively.

GPS TEC values at 22h00 UT, 04h00 UT and 10h00 UT for Cape Town were also predicted and results compared with the corresponding TEC values from the IRI model (denoted as IRI-2001) in fig. 4.9. In fig. 4.9 large TEC variations are observed during autumn and spring equinoxes, while TEC generally decreases in winter. TEC also depletes on the night side of the Earth due to the recombination of free electrons with the ions and thus diurnal variation is the strongest

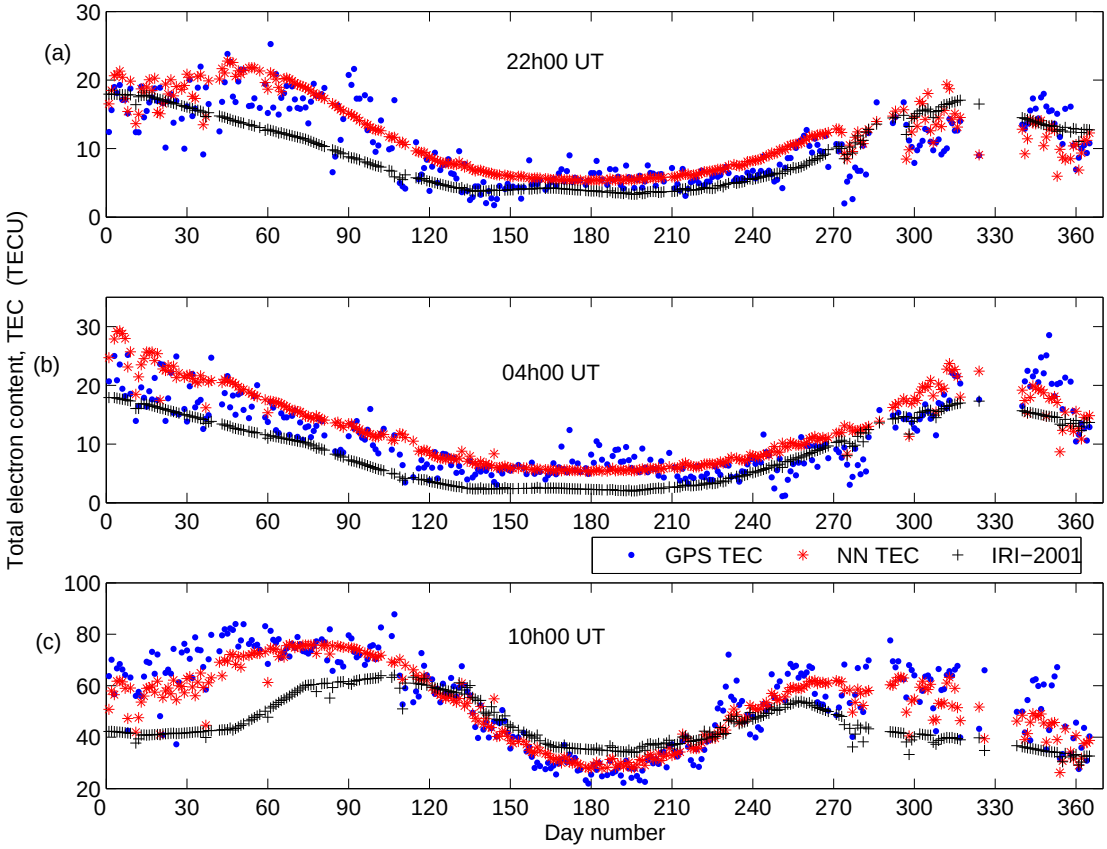


Figure 4.9: Comparison of GPS TEC, NN TEC and IRI TEC values over Cape Town (33.95°S, 18.47°E) in 2002 at (a): midnight (22h00 UT), (b): sunrise (04h00 UT) and (c): midday (10h00 UT).

driving mechanism for TEC changes (Klobuchar, 1991). There is general agreement between GPS TEC and IRI TEC in the shape of seasonal variations as seen in fig. 4.9. A similar seasonal trend was reported by Mosert et al. (2007) using data from the Northern hemisphere station, Ebro (40.8°N, 0.5°E). The significant difference observed, where GPS TEC is greater than IRI TEC, may be due to the altitude at which TEC is estimated in both cases (altitudes are about 20 200 km and 2000 km for GPS TEC and IRI TEC respectively). The differences in TEC values could be compatible with the plasmaspheric electron content present in GPS TEC (e.g. Mosert et al., 2007). However, this can't be the only reason given that IRI TEC contains a modelled component representing the topside ionosphere. Table 4.3 shows the RMSE values between the GPS TEC and TEC values as predicted by the NN and IRI-2001 models for equinox and solstice days. The NN model predicts GPS TEC more accurately for the winter solstice (June 22) than during other seasons. The IRI-2001 model provides a more accurate prediction for the spring equinox (September 23) than the NN model. Table 4.3 shows that the

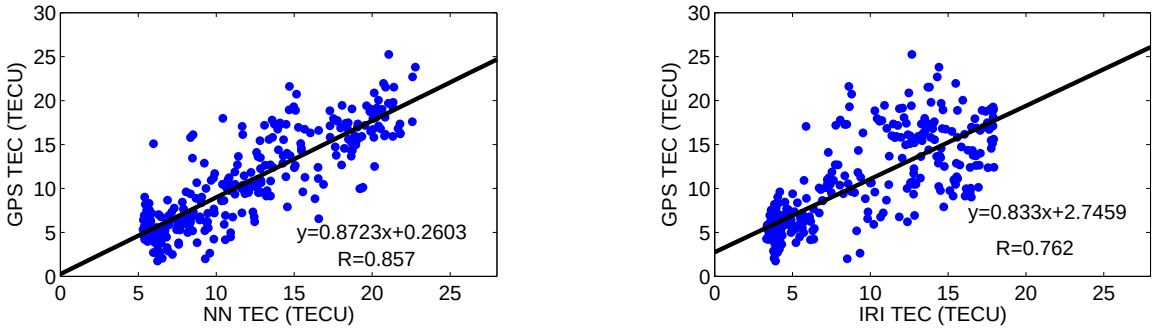
Table 4.3: Computed RMSE values (TECU) for each of the four days representing equinoxes and solstices in 2002 over Cape Town (33.95°S, 18.47°E).

Date of the month in 2002	RMSE (TECU) between GPS TEC	
	NN TEC	IRI TEC
March 20	3.880	15.086
June 22	1.687	4.690
September 23	4.061	3.866
December 21	2.520	2.702

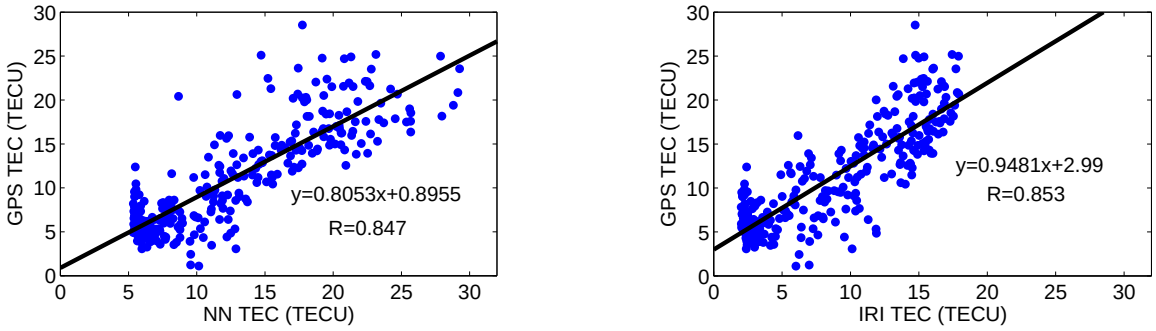
average error difference between the NN TEC and GPS TEC as quantified in terms of RMSE is $\sim$ 3 TECU in contrast to $\sim$ 6 TECU RMSE between the IRI TEC and GPS TEC during equinoxes and solstices. However, taking into account that only one station was analysed here, more work should be done to confirm this result. The overprediction of GPS TEC by the IRI model during winter as shown in fig. 4.8(c) and fig. 4.9(c) may be due to the historical fact that there is generally less data available over the Southern African region (McKinnell, 2002). Very recently Moeketsi et al. (2007) compared GPS TEC derived using the University of New Brunswick (UNB) ionospheric mapping technique (UNB-IMT) with IRI TEC from IRI 2001, over a mid-latitude GPS receiver station located at Sutherland (32.38°S, 20.81°E), South Africa and also obtained an overprediction ($\sim$ 05h00 UT - 10h00 UT) for day 103 in 2003. At the equinoxes, maximum TEC values of $\sim$ 80 TECU and 60 TECU occur at 12h00 UT for March 20 and September 23, 2002, respectively. On the other hand, maximum TEC values of $\sim$ 33 TECU and 34 TECU during the solstices occured at 13h00 UT and 09h00 UT for June 22 and December 21, 2002 respectively. In all cases (equinoxes and solstices) a gradual morning rise and evening decrease of TEC is observed and generally, maximum TEC occurs between 10h00 UT and 12h00 UT. Fig. 4.10 shows the scatter plot for GPS TEC values with corresponding TEC predictions from the NN and IRI-2001 models over Cape Town (33.95°S, 18.47°E) in 2002, with lines of best fit inserted for all cases. Correlation coefficients provide reliability levels of the NN and IRI models' predicted GPS TEC (a correlation coefficient of 1 indicates a perfect fit). Thus, GPS TEC is highly correlated to NN TEC at 10h00 UT with a correlation coefficient (R) of 0.905. The high average plasmaspheric contribution of $\sim$ 4.44 TECU is also observed at 10h00 UT with reference to IRI TEC as shown in fig. 4.10(c). Table 4.4 shows the computed average RMSE values and differences in TEC (ΔTEC) between GPS TEC, IRI TEC and NN TEC for 2002 at 22h00 UT, 04h00 UT and 10h00 UT. The average ΔTEC is defined as:

$$\Delta T_{av} = \frac{1}{N} \sum_{i=1}^{N} (T_i - (T_{j=1,2})_i) \tag{4.10}$$

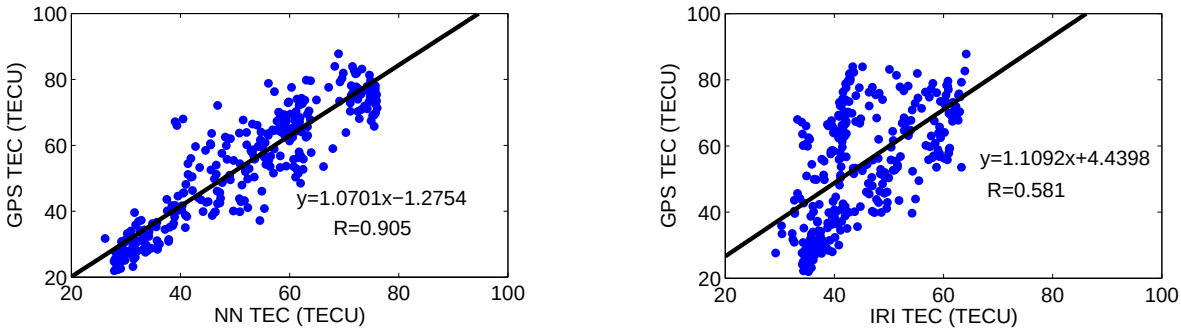
where



(a) A plot of GPS TEC versus NN TEC and IRI TEC at 22h00 UT.



(b) A plot of GPS TEC versus NN TEC and IRI TEC at 04h00 UT.



(c) A plot of GPS TEC versus NN TEC and IRI TEC at 10h00 UT.

Figure 4.10: A scatter plot of GPS TEC versus NN TEC and IRI TEC values over Cape Town (33.95°S, 18.47°E) in 2002 at local midnight, morning sunrise and midday respectively showing correlation coefficients obtained from fitted linear regressions.

ΔT_{av} is the average difference between the GPS TEC (T) and NN TEC ($T_{j=1}$) or IRI TEC ($T_{j=2}$), $i = 1, 2, ..., N$, and N is the total number of observations at a particular time in 2002. The least bias values or systematic errors of estimation show that the NN model predicts GPS TEC more accurately than the IRI-2001 for the considered times. However, the NN model overpredicts GPS TEC during equinoxes in 2002 at 22h00 UT and 04h00 UT.

Table 4.4: Average RMSE values and biases between GPS TEC and predicted values (NN TEC and IRI TEC) at different times in 2002 over Cape Town (33.95°S, 18.47°E).

Time (UT)	RMSE (TECU)		Bias (TECU)	
	NN TEC	IRI TEC	NN TEC	IRI TEC
22h00	3.096	3.768	-1.227	1.207
04h00	3.711	4.023	-1.543	2.553
10h00	7.723	16.825	2.374	9.363

4.7.2 Springbok and Upington

The hourly NN model, which was specifically developed for Springbok from the data available for three years (2002-2004,) was used to predict GPS TEC for Upington during equinoxes and solstices. The measured GPS TEC values were compared with NN TEC, IRI TEC and ionosonde TEC (ITEC) from a South African ionosonde station located at Louisvale (28.5°S, 21.2°E), as shown in fig. 4.11. Although their geographical positions have not been considered so far, the NN optimum architectures at Sutherland and Springbok turned out to be identical (Habarulema et al., 2007). Optimum architecture simply means that the number of hidden nodes that gave the least RMSE and hence better predictions was the same for both stations when considered independently.

With limited data available to develop a fully representative NN model it nevertheless correctly predicts the diurnal variations for equinox and solstice days, as seen in fig. 4.11. Due to the very limited number of GPS data points for June 21 and September 23, winter solstice and spring equinox are represented by June 22 and September 22. Predictions for autumn equinox and summer solstice were obtained for March 19 and January 2 instead of March 21 and December 22 in 2003 due to the unavailability of GPS data at the time this work was done.

Table 4.5: Coefficients of determination (R^2) between GPS TEC and predicted values (NN TEC and IRI TEC) for the four days representing equinoxes and solstices in 2003 over Upington (28.41°S, 21.26°E).

Month date (in 2003)	Time (UT) for the available data	R^2 between GPS TEC and NN TEC	IRI TEC
March 19	04h00 - 17h00	0.9703	0.9287
June 22	04h00 - 17h00	0.9431	0.9409
September 22	05h00 - 17h00	0.9211	0.9456
January 2	04h00 - 16h00	0.9470	0.8452

The Louisvale ionograms for equinox and solstice days in 2003 were manually edited using

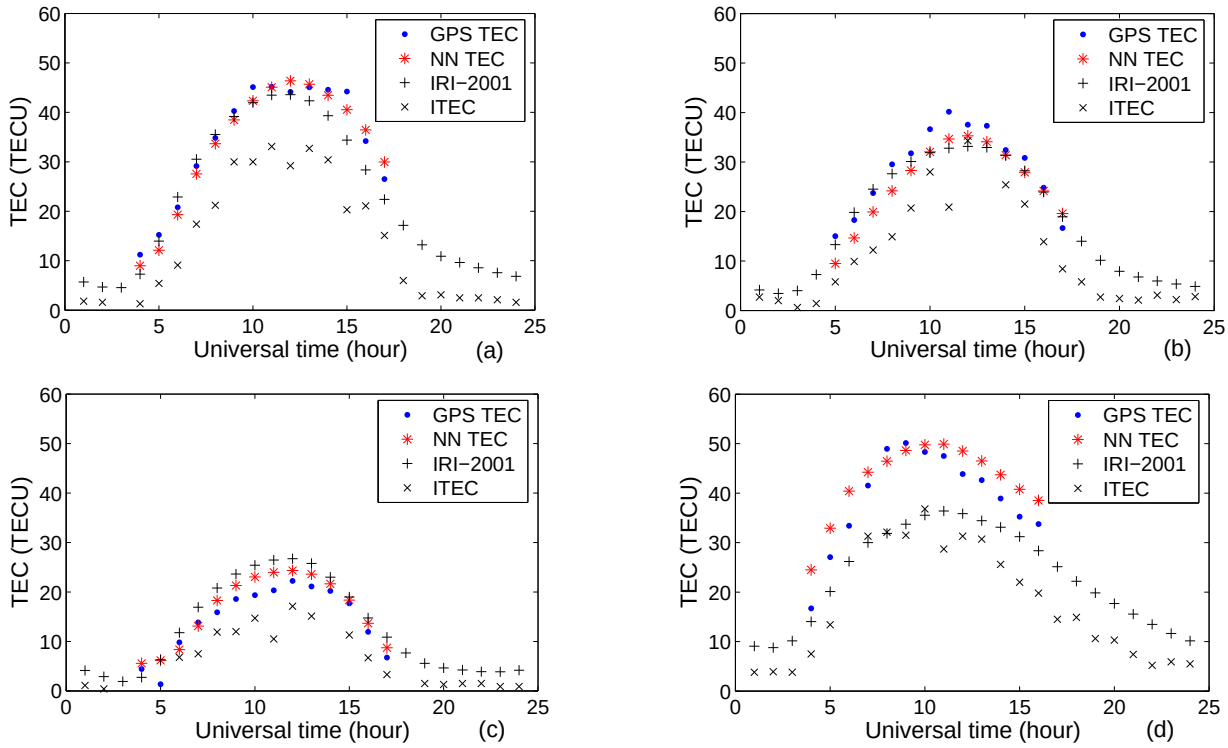


Figure 4.11: TEC variations over Upington during equinoxes: (a) March 19, (b) September 22 and solstices: (c) June 22, (d) January 2; all in 2003.

SAO explorer (Reinisch et al., 2004), from which ITEC values were extracted that represent TEC to an altitude of about 1000 km. The ITEC values were derived from a combination of the measured bottomside and the modelled topside ionosonde profiles. More details about the verification and editing of ionograms obtained from digital pulse sounders can be found in Reinisch et al. (2004). The IRI overpredicts ITEC as expected for the four days representing equinoxes and solstices in 2003. This is due to the fact that the IRI model estimates TEC values up to an altitude of 2000 km (e.g. Bilitza, 2001), and that the topside electron density profile contributes a large percentage of TEC (e.g. Bilitza, 2001; Mosert et al., 2007). Hence, the difference between ITEC and IRI TEC is as a result of the IRI topside profile, which is an analytical representation of the Bent model (e.g. Bilitza et al., 2006; Mosert et al., 2007). It is however important to note that the ionosonde used is not exactly co-located with the GPS receiver, and the difference in location could also contribute to the differences seen in the TEC values.

From the available data, the NN predicts GPS TEC more accurately during the autumn equinox, while the IRI achieves its best prediction during the spring equinox. Table 4.5 shows a summary of coefficients of determination (R^2) between GPS TEC and predicted values (NN TEC and IRI TEC) for the four days representing equinoxes and solstices in 2003 over Up-

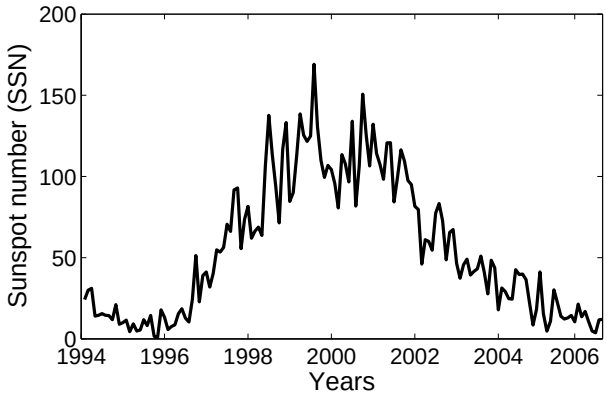


Figure 4.12: Monthly averages of sunspot numbers from 1995 to June 2007. Data obtained from http://solarscience.msfc.nasa.gov/.

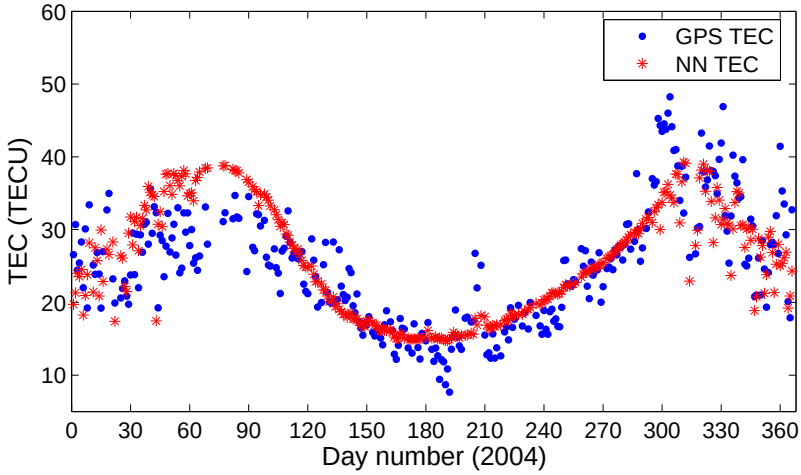
ington (28.41°S, 21.26°E) during the periods when data was available. Although there is a difference in the magnitude of TEC values, it is a common observation for both Cape Town and Upington that the IRI overpredicts GPS TEC during the winter solstice as shown in figs. 4.8(c) and 4.11(c). Comparison of figs. 4.8 and 4.11 shows that TEC is greater in 2002 than in 2003. While this may be an illustration of the latitudinal and longitudinal dependance of TEC, since there are two different GPS stations under consideration, TEC also has a strong solar activity dependance and fig. 4.12 shows that 2002 was close to solar maximum, whereas 2003 fell in the declining phase of solar cycle 23.

4.7.3 Sutherland and Springbok

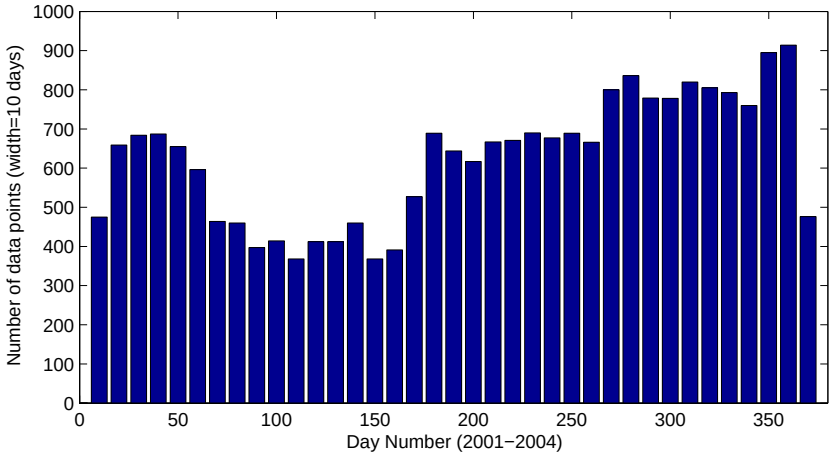
To complete the network of the selected GPS stations in this feasibility study, results from the NN model constructed for Sutherland were used to predict local midday (SAST) TEC values for Springbok in 2004 and results are shown in fig. 4.13(a). This figure shows results for the feasibility study done to determine the suitability of NNs for TEC predictions over South Africa (Habarulema, 2007). It is important to note the correct seasonal representation of TEC by the NN model at a geographic latitudinal difference of $\sim 3^{\circ}$ ($\sim$ 300 km separation), a factor that will later be used to generate simulated TEC values for South Africa as a whole (chapters 6 and 7).

From this figure it is clear that the NN overpredicts GPS TEC, reaching a maximum of $\sim$ 38 TECU in autumn. Four reasons may be put forward to try and explain this:

1. It could be that TEC varies greatly during equinoxes (Kouris et al., 1999) and thus all parameters which influence TEC have high variation levels during this period. This is a factor that is less likely to be the cause in this particular case since TEC variability is higher during the spring equinox.



(a) GPS TEC variation for Springbok (29.67°S, 17.88°E) at 10h00 UT in 2004 along with predicted TEC values generated by the NN hourly model for Sutherland (32.38°S, 20.81°E).



(b) A histogram showing the number of data points used in the NN training during the construction of the hourly model over Sutherland (32.38°S, 20.81°E).

Figure 4.13: Comparison of GPS TEC and NN TEC over Springbok at 10h00 UT in 2004 (top figure) and the histogram showing available data used in NN training in the construction of hourly model over Sutherland (Habarulema, 2007).

2. The unavailability of sufficient data for the period used in NN training prohibited a truly representative input space. "Enough" historic data is a basic requirement for the NN to learn and generalise the non-linear property of any parameter (e.g. Haykin, 1994).

3. The NN model developed from Sutherland data was used to predict TEC at Springbok, however their geographical locations were not included in NN training.

Fig. 4.13(b) is a histogram showing the data points that were used in NN training and validation processes for Sutherland's hourly NN model. There was limited data available during the March equinox compared to other periods and, this is likely to be the cause of the overprediction. The

very low number of data points corresponds to a limited number of all the input parameters which influence TEC during this period. It is a well-established concept that a large database is one of the most important requirements in applications involving NN modelling and therefore NNs may not perform well without enough historic data (Bishop, 1995; Macpherson et al., 1995; Williscroft and Poole, 1996; Conway et al., 1998; McKinnell, 2002). Possibilities of representing TEC variability at high altitudes have been investigated and details will be discussed in chapter 5. Although the sunspot number has been used to represent solar activity, the sun follows a sunspot cycle of $\sim$ 11 years (fig. 4.12) and with a dataset of at most 5 years included in the NN training, the solar variations may not be effectively predicted at this stage.

4.8 Summary and conclusion

In this chapter possibilities of TEC modelling using regression analysis methods and NNs were investigated. It was shown that NNs can more effectively predict GPS TEC values than regression methods considered at locations for which data was not necessarily included in the development of NN and regression models. A comparison was made of GPS TEC (calculated using the ASHA algorithm) and predictions from the IRI-2001 and NN models over a mid-latitude station located at Cape Town (33.95^{o}S, 18.47^{o}E). Results show that on average the NN predicts GPS TEC more accurately than the IRI-2001 model. The overprediction of GPS TEC by the IRI-2001 model is consistent with the findings of Moeketsi et al. (2007), who made a comparison of GPS TEC and IRI TEC over Sutherland (32.38^{o}S, 20.81^{o}E). This has been attributed to a paucity of Southern Africa ionospheric data in the IRI model (McKinnell, 2002). A further comparison of GPS TEC with IRI TEC and ITEC revealed that ITEC is less than its counterparts. This is expected due to the altitudes used to determine these three values. ITEC is determined at best to about 1000 km and the accuracy is determined by the derived electron density profile which consists of a modelled topside. Therefore, it is expected that ITEC will contain some error although it has not been possible to quantify it yet. These similar differences between GPS TEC, IRI TEC and ITEC have been reported in other recent sources (e.g. Mosert et al., 2007; McKinnell et al., 2007; Habarulema et al., 2007). Computation of correlation coefficients and systematic errors of estimation showed that results from the NN model are more reliable than the IRI-2001 model over South African locations with reference to GPS TEC prediction. The NN model achieved its maximum prediction accuracy at winter solstice (June 22) and local midday (10h00 UT) in 2002 for Cape Town observations, with a RMSE of $\sim$ 1.69 TECU and a correlation coefficient of 0.905. The confirmation of the NN accuracy in predicting GPS TEC at Cape Town is also reflected in the average systematic errors of estimation as shown in table 4.4. On the other hand, the IRI-2001 model predicted GPS TEC more accurately than the NN model at spring equinox with an average RMSE of $\sim$

3.87 TECU (table 4.3) and correlation coefficient (R) of ~ 0.97 (table 4.5) at Cape Town and Upington respectively. It has not been possible to make concrete conclusions from the GPS TEC predictions obtained over Upington in 2003, due to a very limited database used in the NN training and the incompleteness of the available diurnal GPS data during equinoxes and solstices. However the available data shows that the NN reflects the TEC diurnal shape and the predicted NN TEC shows a strong correlation with GPS TEC. This will be pursued in future as more GPS data becomes available. Only five GPS locations out of a total of about 50 (by $\sim$ March 2009) possible locations currently present in South Africa, have been used for the preliminary studies. Also not included in this study is the geographical location representation since it only dealt with the development of single station models. This parameter is considered in chapters 6 and 7 which also include the prediction of GPS TEC (derived using the ASHA algorithm) and comparison with the IRI-2007 model at other locations within South Africa, especially at GPS stations that are in close proximity to other ionosonde stations, for validation and investigation of the latitudinal and longitudinal dependence of TEC in NN modelling over the South African GPS receiver stations.

Chapter 5

The solar wind and TEC variations

This chapter attempts to investigate the statistical relationship between selected solar wind parameters and GPS-derived TEC by building on the acquired knowledge in chapter 4. In a similar way, the model developed for a particular GPS station is used to reproduce modelled TEC values for a different station, but, in this case, with the solar wind parameters included. Another significant difference is that recurrent neural networks, specifically Elman networks (ENNs), have been introduced in TEC modelling at this stage as a proxy for time delay correction within the modelling process, in an effort to "synchronise" solar wind satellite data with GPS TEC data.

5.1 Introduction

The solar wind - magnetosphere - ionosphere (SW-MAG-ION) system is not an easy phenomenon to break down, so as to deal with the respective physical mechanisms responsible for the behaviour of different constituent ions and molecules that contribute to the changes exhibited by ionospheric parameters. According to Lyon (2000), the SW-MAG-ION form a single system that owes its existence to the energy and momentum transfer from the solar wind to the ionosphere through the magnetosphere, a concept that can be summarised as "solar wind - magnetosphere - ionosphere coupling". Among the ionospheric parameters influenced by this transfer of energy is TEC, especially during geomagnetic storms. Preliminary studies indicate that South Africa's ionospheric TEC behaviour is characterised by both negative and positive storm effects during disturbed conditions (Habarulema et al., 2009c), although the physical mechanisms responsible for TEC behaviour are generally not well understood due to the limited number of studies that have been carried out for this region. In this chapter the first four standard parameters that influence TEC variability (as described in chapter 4) are combined with each selected solar wind parameter in an attempt to characterise the quantitative behaviour of TEC with respect to changes in solar wind dynamics. The solar wind param-

eters under investigation are the B_z component of the interplanetary magnetic field (IMF), proton number density (N_p), solar wind bulk velocity (V_{sw}) and the solar wind dynamic pressure (P_{dyn}). The B_z component is considered in empirical modelling due to the fact that its southward turning is an indication that a magnetic storm is in progress (e.g. Tsurutani and Gonzalez, 1997), which later influences the ionospheric TEC negatively and/or positively. The P_{dyn}, which primarily controls the position of the magnetosphere, is directly related to N_p and V_{sw} ($P_{dyn} \propto N_p V_{sw}^2$) making these three possible candidates to represent solar wind changes and its influence on ionospheric behaviour during statistical modelling. Due to the complexity of the entire system, separating individual agents and quantifying their respective contributions to ionospheric TEC variability is a complex problem. As a result, a non-linear approach was employed to extract statistical information about TEC variations due to solar wind parameter fluctuations.

5.2 Solar wind data

By definition the solar wind is a stream of energetic particles that constantly flows from the sun and reaches the Earth's atmosphere at a typical speed within the range of $\sim$ 300 - 700 km/s. The solar wind consists of about 95 % electrons and protons and they are almost equal in proportion. Solar wind data has been gathered by various spacecrafts, such as the Advanced Composition Explorer (ACE), SOHO, Wind, Geotail, etc, and is publically and readily available. Due to data gaps and "bad" measurements that exist in some datasets as a result of instrument failures and spacecrafts probably straying from their orbits, efforts have been undertaken to consolidate the available data and make necessary corrections (e.g. King and Papitashvilli, 2003). This gave rise to the preparation of comprehensive databases hosted and maintained by the National Space Science Data Center (NSSDC) and the NASA Goddard Space Flight Center and its affiliations. This data archiving has evolved from old OMNI, OMNI 1 to the present OMNI 2 datasets. In this chapter both corrected (for time delay) and uncorrected data (from OMNI 2 and ACE satellite datasets respectively) are used to attempt a comparative analysis.

5.2.1 ACE data

The ACE spacecraft project was developed by the John Hopkins University Applied Physics Laboratory (JHU/APL) and launched on 25 August 1997, with the primary objective of providing data for the study of the elemental and isotopic composition of the solar corona, interplanetary medium, galactic matter and local interstellar medium, among others (Stone et al., 1998). It has a total of one engineering and 9 scientific instruments on board. Relevant to this study are the Solar Wind Electron, Proton and Alpha Monitor (SWEPAM) and the Magnetic Field Experiment (MAG). SWEPAM measures the proton (H, H_e) charge distributions

as well as the electron energy of the incoming solar wind from which data for V_{sw} and N_p are derived, while MAG provides the x, y and z components of the IMF in geocentric solar magnetospheric (GSM) coordinates (Stone et al., 1998; Chiu et al., 1998). Data for these parameters are available online and were downloaded from http://cdaweb.gsfc.nasa.gov/cdaweb/sp_phys/ (GSFC, 2007). The ACE spacecraft orbits around the L1 Sun-Earth Libration point that is approximately 1.5×10^6 km ($230R_E$) from the Earth along the Sun-Earth line. The Sun, Earth and ACE make a three-body system, with the latter appearing stationary with respect to the first two, and thus the position of the observer is approximately constant, thereby eliminating problems related to the observer's motion. At the L1 point, ACE is in a favourable position to record measurements of the IMF and solar wind properties before the whole process impacts on the magnetosphere which is a key requirement for any spacecraft to provide continuous information on solar energy and interplanetary particles, the solar wind and cosmic rays (Stone et al., 1998). More details about the ACE spacecraft are found in Stone et al. (1998) and Chiu et al. (1998).
In the context of this study, when all other parameters that influence TEC variability are fixed and the ACE solar wind data is included in the empirical model, one could possibly conclude that any variations in results are due to the solar wind contribution through magnetosphere-ionosphere coupling.

5.2.2 OMNI 2 data

The OMNI dataset consists of low and high resolution data for IMF and solar wind plasma parameters, as measured by various spacecrafts near the Earth's orbit (GSFC, 2009). Spacecrafts that contribute data to the OMNI dataset include ACE, Wind, ISEE 3, Geotail and IMP 8 (e.g. King and Papitashvilli, 2003). However, the IMP 8 magnetometer failed in June 2000 (King and Papitashvilli, 2003) and only a fraction of its data may have been included in the dataset (2000-2007) of interest in this chapter. While these satellites are orbiting at different altitudes with respect to each other, the OMNI data is corrected for time delay (time shifted) and is generally treated as observational data made from the bow shock nose. For both ACE and OMNI 2 datasets, the P_{dyn} is derived from N_p and V_{sw} data using the relation

$$P_{dyn} \propto N_p V_{sw}^2 \tag{5.1}$$

P_{dyn}: solar wind dynamic (flow) pressure (nPa),
N_p: proton number density (N/cm^3),
V_{sw}: solar wind bulk speed (velocity) in km/s.

The ACE satellite dataset can be considered a "subset" of the OMNI 2 data since the OMNI data is derived from different spacecrafts' datasets, including ACE. Fig. 5.1 shows an example

of the comparison of V_{sw} obtained from ACE and OMNI datasets along with the computed ACE delay time in minutes. The OMNI dataset V_{sw} appears to be a "fit" for the ACE dataset

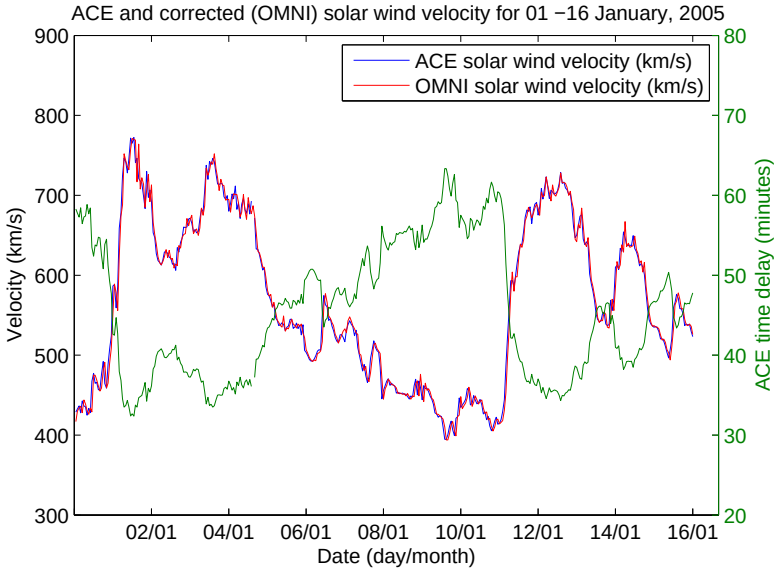


Figure 5.1: Hourly solar wind velocity (km/s) measured by ACE and corrected OMNI data during 2005 with the delay time (computed by dividing 1.5×10^6 km by the ACE solar wind velocity) from the L1 libration point to the Earth.

velocity and vice versa. The two datasets do not differ significantly, mainly because of the low resolution (hourly) data used. The ACE time delay (green plot in fig. 5.1) is highly variable, a factor which motivates the use of recurrent networks as will be discussed in sections 5.4 and 5.5.

5.2.3 Data compensation

The availability of GPS TEC data was taken as the reference for the consideration of solar wind data. The GPS TEC database contained periods with data, but, in some cases, without V_{sw}, N_p, IMF B_z or P_{dyn} data. These data gaps in the ACE and OMNI data were filled with monthly hourly average values for the periods where data existed within a particular month, according to the following equation:

$$Y_m = \frac{1}{n_d} \sum_{i=1}^{n_d} X_{j=1,2,3,4} \tag{5.2}$$

Y_m: missing data in V_{sw} ($X_{j=1}$), N_p ($X_{j=2}$), Bz ($X_{j=3}$) and P_{dyn} ($X_{j=4}$),
i=1, 2,..., n_d, with n_d being the total number of hourly data points for which actual data is available in a given month.

The most significant amount of missing data was found in the N_p parameter, followed by V_{sw} and finally B_z. There is no ideal method of data compensation. Therefore, monthly mean values were taken as proxies for missing data only within the solar wind database, simply because the removal of GPS data for periods where solar wind data was nonexistent would eliminate useful information about TEC variability which is the primary focus of this study. Also the technique used (NNs) extracts average statistical information and tends to ignore the "extreme" values unless they are taken into account during model development.

5.3 Relationship between solar wind and TEC

As already mentioned, the SW-MAG-ION coupling process is mainly driven by the energy and momentum released by the solar wind. The interaction between the solar wind and the magnetosphere depends mainly on the magnetic field in the solar wind. Due to the existance of magnetosphere-ionosphere coupling, changes in the magnetosphere result in corresponding ionospheric variation. According to Shue et al. (1998), the magnetosphere's size is controlled by the P_{dyn}. During extreme conditions such as magnetic storms, the P_{dyn} increases, thereby compressing the magnetosphere. Huang et al. (2005) and Huang and Foster (2001) explain the direct relationship between the solar wind and ionospheric parameters (e.g. F-peak altitude, hmF2 and maximum electron density, NmF2) over mid-latitude locations using data collected from a location in the Northern hemisphere, namely Millstone Hill radar (42.6°N, 71.5°W). These authors suggest different mechanisms to explain the sudden decrease of ionospheric plasma density as a result of changes in either solar wind or IMF, if not both. One suggestion is based on the basic physics principle that magnetosphere expansion caused by the sudden decrease of P_{dyn}, increases the volume of the plasmasphere, resulting in lower plasma density and plasma pressure. Consequently, there is an upward flow of plasma which reduces the ionospheric electron density until the process of plasmaspheric refilling begins which restores the pressure balance. This suggestion however does not take into account the recombination and ionisation effects.

5.3.1 Plasmaspheric TEC and solar wind parameters

The plasmasphere is a region above the topside ionosphere. Its lower boundary is believed to be the transition height where O^+ ions become less dominant than H^+ and He^+ ions (Belehaki et al., 2004). The plasmasphere consists of low energy particles. The slow particle motion is dominated mainly by the geomagnetic field. In this chapter the plasmasphere is defined as a region above 1000 km, since ionosonde TEC will be taken as a benchmark when determining plasmaspheric TEC from GPS data (Belehaki et al., 2004; McKinnell et al., 2007). The solar wind parameter is considered in this empirical modelling with the prospect of capturing

plasmapheric TEC variations during magnetic storms, especially at high altitudes. However, a very recent study pointed out that separating agents forming ionospheric disturbances during storm conditions is too difficult due to the complexity of the correlation between solar wind dynamics and the ionospheric variations (Biktash et al., 2008). Furthermore, Biktash et al. (2008) showed that the solar wind plays an important role in controlling the equatorial ionosphere during geomagnetic storms, by explaining the IMF B_z effects on ionospheric parameters such as nighttime equatorial foF2 and virtual heights (h'F). Jakowski et al. (1999) studied the relationship between the solar wind and TEC before and around solar minimum at different European latitudes using V_{sw} as one of the parameters and found that there exists an anticorrelation. The observed delay between V_{sw} and TEC was interpreted as an indication of close coupling of the solar wind with the ionosphere-thermosphere dynamics related to the observed negative phases of ionospheric storms a few days after storm onset. Sethia et al. (1978) also showed that there exists an inverse relationship between plasmaspheric TEC and V_{sw}, in a study over an equatorial station, Ootacamund (11.4° N, 76.6° E). These authors utilized TEC measurements derived from Faraday rotation and group delay methods, using radio beacon signals from the geostationary Application Technology Satellite (ATS-6) to estimate plasmaspheric TEC.

On the other hand, a decrease of TEC (from both GPS and ionosonde data) has been observed over South Africa during the main phases of geomagnetic storms (Opperman, 2008). During this period, an increase in V_{sw}, N_p and subsequently P_{dyn} was observed. An analysis during storm conditions was done for a Southern hemisphere mid-latitude station located at Grahamstown (33.30° S, 26.53° E), South Africa, where the ionosonde and GPS receiver are co-located. Plasmaspheric TEC is simply defined as

$$\text{PTEC=GTEC- ITEC} \tag{5.3}$$

PTEC, GTEC and ITEC are plasmaspheric TEC, GPS TEC and ionosonde TEC respectively.

Fig. 5.2 shows the variation of the average PTEC with hourly average values of IMF B_z (nT), N_p (cm^{-3}), V_{sw} (km/s) and P_{dyn} (nPa) over Grahamstown for 5-18 May 2005 with their respective linear relations given approximately as;

$$\left.\begin{array}{ll} \text{(i)} & B_z = 0.9316P_T - 3.119, \\ \text{(ii)} & N_p = 0.7056P_T + 3.7625, \\ \text{(iii)} & V_{sw} = 63.179P_T + 334.1 \\ \text{(iv)} & P_{dyn} = 0.936P_T - 0.8119 \end{array}\right\} \tag{5.4}$$

where P_T represents PTEC in TECU. Statistical computation of correlation coefficients (μ) between PTEC and solar wind parameters gives $\sim$ 0.573, 0.198, 0.656 and 0.772 for B_z, N_p, V_{sw} and P_{dyn} respectively.

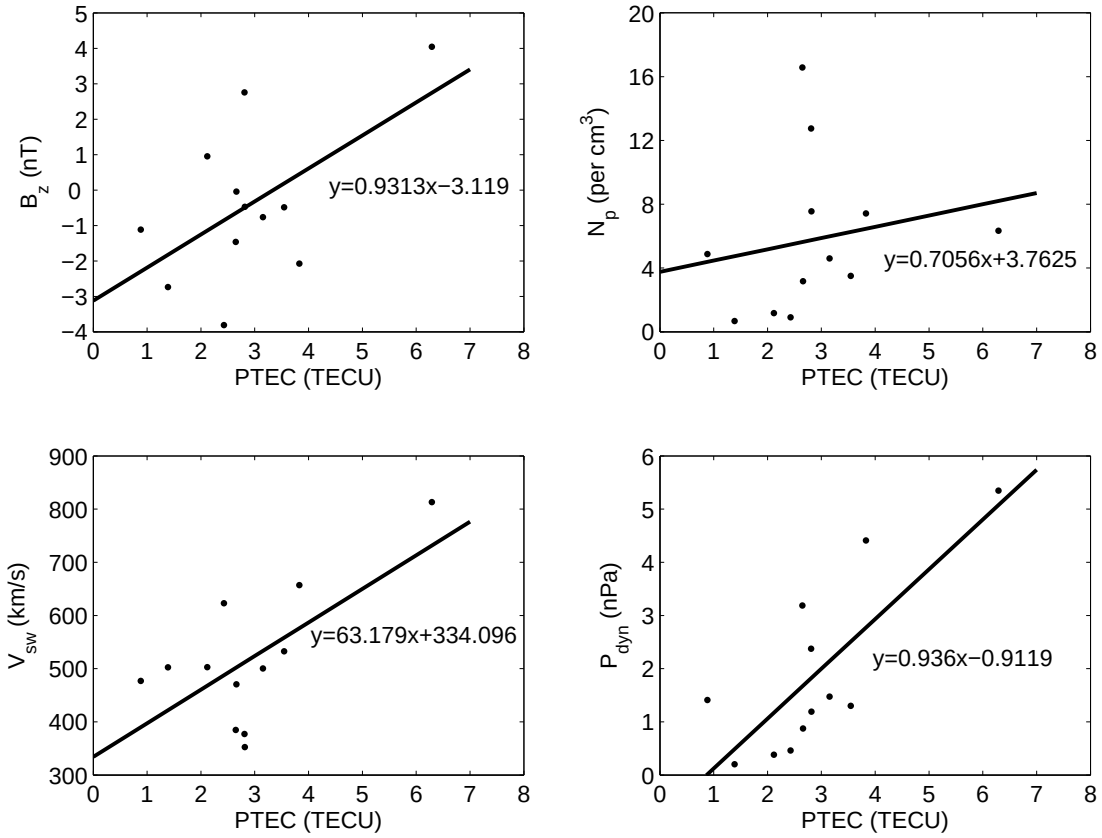


Figure 5.2: Variation of average values for plasmaspheric TEC with solar wind parameters (data from ACE) during 5-18 May 2005.

The plasmaspheric content increased for the storm day (15 May) reaching ~ 6 TECU while it fell within $\sim 2-3$ TECU during the relatively quiet days before and after the storm. Belehaki et al. (2004) also reported an increase in plasmaspheric electron content on the day of storm initiation and a drop in PTEC a few days after the storm giving clear evidence of plasmasphere depletion. The first comparative study between GTEC and ITEC over Grahamstown for the period of March-June 2005 revealed that on average, the latter is less than the former by $\sim$ 3.57 TECU (McKinnell et al., 2007). In the context of relating the solar wind parameters to plasmaspheric electron content, significant correlations were obtained for B_z, V_{sw} and P_{dyn} with N_p as the least correlated parameter to PTEC for the analysed period.

5.4 Recurrent networks in space weather modelling

As explained in chapter 2, NNs are powerful tools that can be used to perform the tasks of learning and generalising the variation of behaviour patterns of parameters that exhibit non-linear characteristics, by means of the input-output mapping process. Different predictions require different inputs, depending on the nature and complexity level of the problems in space weather monitoring, especially where solar wind data is involved. Weigel et al. (2002, 2003) used recurrent NNs to map solar wind measurements with respect to the rate of change of geomagnetic field levels and their predictability. These authors demonstrated that the solar wind and rate of change of the geomagnetic field are related and have a strong dependence on local time and latitude. Vandegriff et al. (2005) fed multiple inputs (including proton intensities with varying energies) into ENNs to demonstrate that they (recurrent NNs) were capable of predicting the arrival of interplanetary shocks on the Earth. Marra and Marabito (2005) explored the ENN technique for solar activity forecasting by considering the sunspot-related time series. Their work is an example of using a single input to predict one output. A comprehensive study to demonstrate the solar wind-magnetosphere coupling by means of ENN modelling using solar wind plasma, IMF and Dst data, was done by Wu and Lundstedt (1997). All these studies show that NNs have various ways and algorithms available for their implementation, such as the Levenberg-Marquardt algorithm which has the advantage that it conserves time, and the back propagation algorithm which is simple to use in both feed forward and recurrent networks (Haykin, 1994; Bishop, 1995).

As stated in chapter 2, this chapter uses one form of recurrent networks known as the Elman networks, which handles function learning based on a combination of current input parameters at any time and a series of previously learned parameters with the corresponding outputs of the entire network connection (Elman, 1990; Haykin, 1994; Bishop, 1995; Marra and Marabito, 2005). One of the major differences between ordinary feed forward NNs (used in chapter 4) and recurrent ENNs is that the latter has a copy layer of the hidden neurons (Haykin, 1994).

The output for single hidden layered ENN, with n input units, m hidden units and one output unit, can be expressed in terms of the activation function and associated weights along with the corresponding biases at different times in an iterative process (Wu and Lundstedt, 1997; Bodén, 2002) as follows;

$$X_j(t) = A_f\left(\sum_{k=1}^{n} v_{jk}x_k(t) + \sum_{h=1}^{m} u_{jh}y_h(t-1) + b_j\right)$$

$$X_o(t) = \sum_{j=1}^{m} X_j(t)w_{oj} + b_o$$

$X_j(t)$: output of the hidden unit j at time t,
$A_f = 1/(1+e^{-x})$: activation function, x is the input training vector,
j, h: indices for the hidden and context nodes,
k, o: indices for input and output nodes respectively,
$X_o(t)$: output of the output unit o at time t,
v, u, w: connection weights associated with the input, hidden and output layers, respectively and b is the bias.

Details about ENNs are provided in chapter 2.

5.5 Inclusion of the solar wind parameter into the Elman network

As far as is known, no extensive works have been published about the quantification of solar wind effects on GPS TEC over the African sector. The attempt done by Habarulema (2007) on a small database over the Square Kilometer Array (SKA) Hub location (30.71°S, 21.39°E), South Africa at 10h00 UT showed that a parameter representing solar wind could be an additional input for the NN to learn, generalise and reproduce TEC patterns. Habarulema (2007) considered V_{sw} and N_p (since electron number density was not available from the ACE satellite dataset) as separate inputs to the NN and the quantification results by means of the RMSE method were generally inconclusive. However, some literature exists on other regions (especially the Northern hemisphere) which links the solar wind variation to midlatitude ionospheric variation (e.g. Huang and Foster, 2001; Huang et al., 2005; Kumar et al., 2008; Denton et al., 2009). While the previous attempt over the SKA Hub used ordinary feed forward networks, the latter study tried to take into account the time delay that exists within the solar wind data with respect to TEC by feeding the data into recurrent networks. This work is also a more comprehensive extension in a sense that it statistically investigates the variations of IMF B_z and P_{dyn} with TEC fluctuations. The use of Elman networks was necessitated by the underlying and significant delay that exists between the time of the solar wind measurement by ACE at L1 libration point and the time it takes to cause the geomagnetic effects on the magnetosphere-ionosphere system. If for example, this delay were a constant, the correction would be easily

achieved by shifting the solar wind measurement by a constant factor. However, the ACE time delay as shown in fig. 5.1 illustrates that this delay is not constant and, therefore, its high variability and the accurate individual correction of each data point and parameter is very difficult, if not impossible. A technique that inherently provides this correction was, therefore, required. The internal structure of Elman networks corrects for this time delay by means of internal memorization of its previous states and the presentation of each state at time t along with the input training pattern at time $t-1$. It is thus adequate to finally mention that recurrent (Elman) networks were chosen for this specific task, because of their ability to store useful information about past data points by means of the additional copy layer and therefore enhances their prediction capability.

5.6 Data and Elman network model setup

5.6.1 Data

Hourly GPS TEC data for the GPS sites located at Sutherland and Cape Town, South Africa (equipped with Ashtech geodetic grade dual frequency receivers) were estimated using the ASHA algorithm. Chapters 2 and 4 describe the ionospheric variability as influenced by seasonal and diurnal variations, solar and magnetic activities and the geographic position of the GPS receiver. Only the first four parameters are considered since this study involves single station modelling. Clearly, a NN would not learn anything about TEC variability with respect to a constant geographical position. The solar and magnetic activity representations determined in chapter 4 were adopted as well as the sine and cosine component methods for day number (D_n) and hour (H_r) in order to take into account the seasonal and diurnal variability of TEC behaviour. The dataset under consideration, consisting of $D_n, H_r, R4, A8$ and the solar wind parameters, were divided into training and validation patterns according to the ratio 7:3, respectively. Hourly data for 2000-2007 over Sutherland were used for the training and validation processes while hourly data for 2005 over Cape Town were used for verification. Fig. 5.3 shows the GPS TEC data that was used for model development, along with daily sunspot numbers to demonstrate the correlation between solar activity and ionospheric TEC behaviour. Overall, the data employed in model development consisted of 51247 (out of possible 67206) data points, each comprising $D_n, H_r, R4, A8$ and either B_z, V_{sw}, N_p or P_{dyn}. Certainly, the verification dataset was not included in training (since it is also from a different location), but the time period chosen for the verification station fell within the training period, thus taking into account the solar activity level changes, as clearly indicated in fig. 5.3. In chapter 4 it was shown that a NN model developed for a particular GPS station can be used to predict TEC variability for a GPS receiver station, the data of which was not involved in the NN model development. This has demonstrated the temporal, and to some extent, spatial capability of

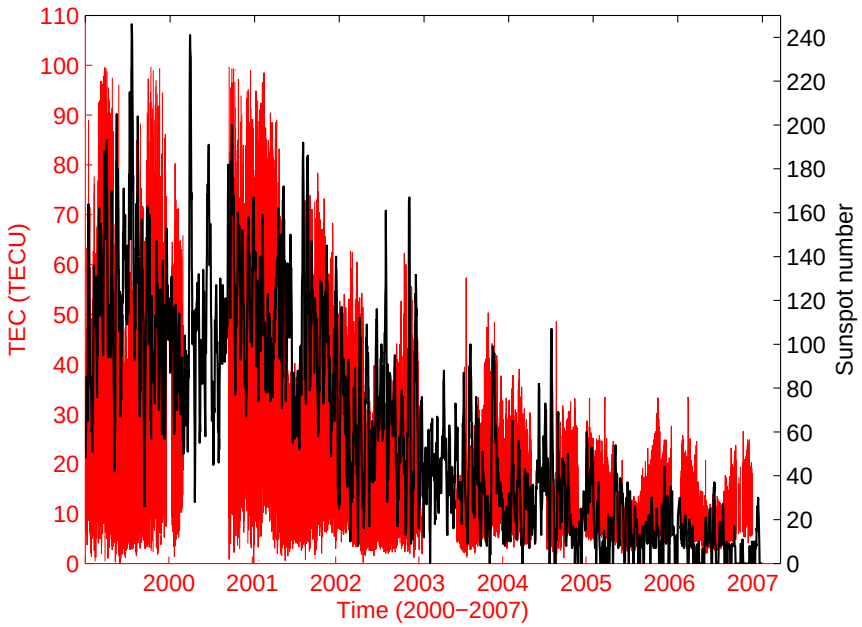


Figure 5.3: Hourly GPS TEC data available over Sutherland along with daily sunspot number for 2000-2007.

the NN in dealing with such non-linear approximations within a particular latitudinal range (Habarulema et al., 2009a). The ability of the NN model to capture TEC variability for station with data from that station not included in training, may be attributed to the fact that, since the stations have a latitudinal difference of $\sim 2°$, the ionospheric pierce point (IPP) which provides an input for the GPS vertical TEC on the receiver used in the NN training (Sutherland), may cover the geographical location of the receiver located at Cape Town.

5.6.2 Elman network model setup

This study was conducted in two phases. The first one involved the prediction of TEC as a function of $D_n, H_r, A8$ and $R4$, which acted as a benchmark while quantifying the solar wind effects on GPS TEC. The second one dealt with the separate inclusion of each solar wind parameter (from ACE and OMNI datasets at different instants), while the parameters in the first option were kept constant. Two network setups were constructed, trained and verified for almost the entire year (2005), which fell in the declining phase of SC 23. It is therefore imperative to mention that the conclusions derived from the predictions do not necessarily reflect the outcomes for high solar activity periods since the verification dataset fell within the low solar activity period. A number of recurrent ENNs were trained in a search for the optimum architecture. Training was monitored until the mean square errors on the testing dataset decreased (Habarulema et al., 2007). There is no direct way of determining the number of

hidden nodes in the hidden layer. Increasing the number of hidden layers does not significantly change the accuracy of results, but makes training easier in terms of time, in some cases (Haykin, 1994). Fig. 5.4 shows a schematic illustration of the Elman network setup. The input training patterns from input nodes $u_{j1}(t)$, $u_{j2}(t)$, . . ., $u_{jm}(t)$ are forwarded to hidden nodes $h_{j1}(t)$, $h_{j2}(t)$, . . ., $h_{jn}(t)$. The outputs of these hidden nodes are propagated back to their respective context nodes $h_{j1}(t-1)$, $h_{j2}(t-1)$, . . ., $h_{jn}(t-1)$ before the computation of the final output $y_j(t)$ for a specified number of iterations.

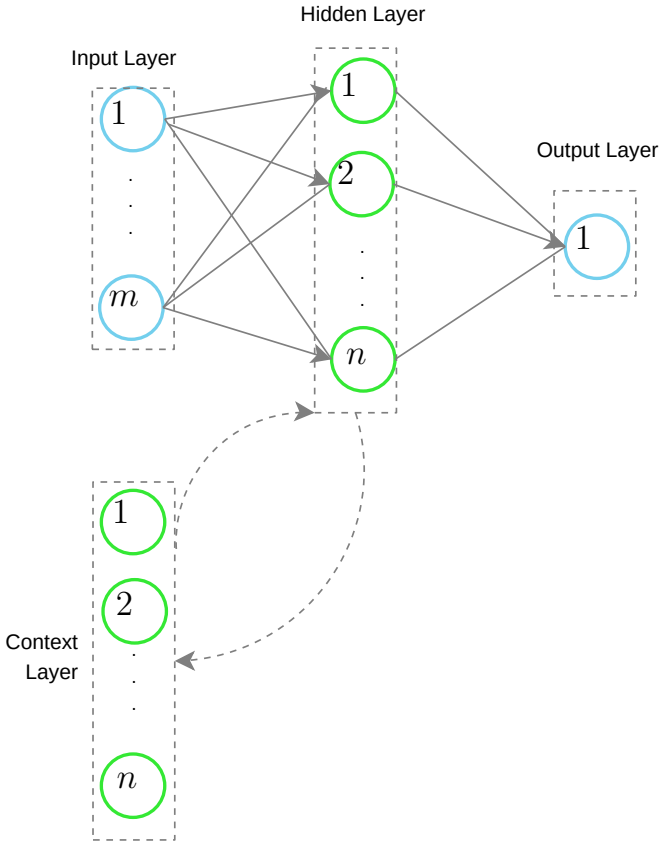


Figure 5.4: A schematic illustration of the Elman network setup. $m = 6, 7$ and $n = 8, 9$ for S_p and after the inclusion of the solar wind parameter respectively.

The function T(S_p) is defined to represent predicted TEC as a function of $D_n, H_r, A8$ and $R4$ i.e. T(S_p)==T($D_n, H_r, A8, R4$) so that T(S_p, B_z), T(S_p, V_{sw}), T(S_p, N_p) and T(S_p, P_{dyn}) represent predicted TEC values after separately adding B_z, V_{sw}, N_p and P_{dyn} respectively to S_p. The architecture consisted of one input layer, one hidden layer and one output layer, the latter having one output neuron. For determination of the optimum NN architecture, the RMSE method was used. The procedure involved the addition of one hidden node at a time, training the NN, testing it with data and finally computing the RMSE between GPS TEC and NN predicted TEC values (Habarulema et al., 2009a). The NN architecture that gave the least RMSE was adopted as the one suitable for TEC prediction. Table 5.1 shows the RMSE values

obtained for different network architectures on different days.

Table 5.1: RMSE (between predicted and measured GPS TEC) for some days in 2005 for different architectures when $D_n, H_r, R4$ and $A8$ are included in the ENN as inputs.

Network denotation	Architecture configuration	RMSE (TECU) between T^m and T^p			
		day 60	day 100	day 250	day 291
A	6:6:1	13.3247	8.3594	5.8260	6.7112
B	6:7:1	2.0740	2.3914	2.0723	2.4126
C	**6:8:1**	**2.1627**	**1.7717**	**1.3522**	**1.9618**
D	6:9:1	2.6417	3.2795	1.8668	1.9108
E	6:10:1	2.8451	2.3629	1.8298	1.9035

The computation of correlation coefficients was used as a complementary option during the NN architecture determination. The correlation coefficient μ is defined as

$$\mu = \frac{1}{\delta_m \delta_p} \left(\sum_{i=1}^{N} (T_i^m - \overline{T}^m)(T_i^p - \overline{T}^p) \right) \tag{5.5}$$

where δ_m and δ_p are the respective standard deviations for measured and predicted TEC defined by

$$\delta_m = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (T_i^m - \overline{T}^m)^2}, \quad \delta_p = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (T_i^p - \overline{T}^p)^2}, \tag{5.6}$$

T^m, T^p: measured and predicted TEC,
$\overline{T}^m$, $\overline{T}^p$: respective means for T^m and T^p over a length interval of the total number of observations N under consideration.

Table 5.1 shows that the network **C** with configuration 6:8:1 was adopted for TEC predictions when standard parameters were incorporated into the ENN as inputs. This configuration was altered to 7:9:1 during the solar wind effects' statistical determination on GPS TEC. Thus the number of hidden nodes that provide the optimum architecture was found to be ∼n+2, where n is the number of inputs to the ENN. The final results, which appear in the following section, were obtained by predicting hourly TEC (T_p) as a function of either six or seven parameters, and may be rewritten as

$$T_p \equiv \begin{cases} f(S_p) & \text{for standard parameters,} \\ f(S_p, V_{sw}) & \text{V}_{sw} \text{ included,} \\ f(S_p, N_p) & \text{N}_p \text{ included,} \\ f(S_p, B_z) & \text{IMF B}_z \text{ included,} \\ f(S_p, P_{dyn}) & \text{P}_{dyn} \text{ included.} \end{cases} \tag{5.7}$$

It must again be stressed that results from the ENN trained for Sutherland (32.38°S, 20.81°E) were used predict/model TEC changes for a receiver site located at Cape Town (33.95°S, 18.47°E). The basis for this is the feasibility study results presented in chapter 4, which demonstrated that it is possible to predict TEC variations for GPS receiver stations where their data was not included in training, within a latitudinal range of $\sim$ 1 - 3 degrees. It will be explained later that this is the proof of NNs' ability to characterise, to a certain degree, the TEC spatial variability.

5.7 Results and discussion

Predictions were obtained for both geomagnetically quiet and disturbed days in order to assess the performance of the ENN model under all conditions. Both diurnal and seasonal TEC modelled results were discussed.

5.7.1 Diurnal TEC predictions

5.7.1.1 During quiet conditions

Figs. 5.5-5.6 show the comparison of hourly GPS TEC and the predicted values as a function of the standard parameters (S_p) and each individual solar wind parameter for days 58 and 285 in 2005 over Cape Town (33.95°S, 18.47°E). The superscripts A and O on B_z, V_{sw}, N_p and P_{dyn} indicate that the solar wind data used in the ENN model development and verification correspond to the ACE and OMNI 2 datasets respectively.

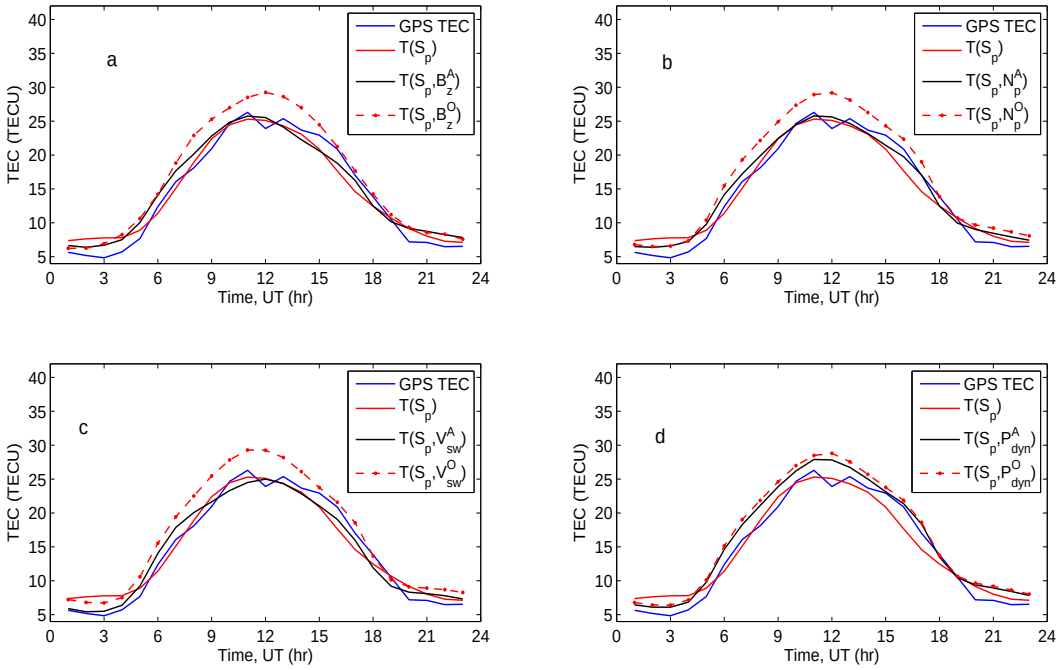


Figure 5.5: Comparison between hourly GPS TEC and the corresponding predictions for day 58 in 2005.

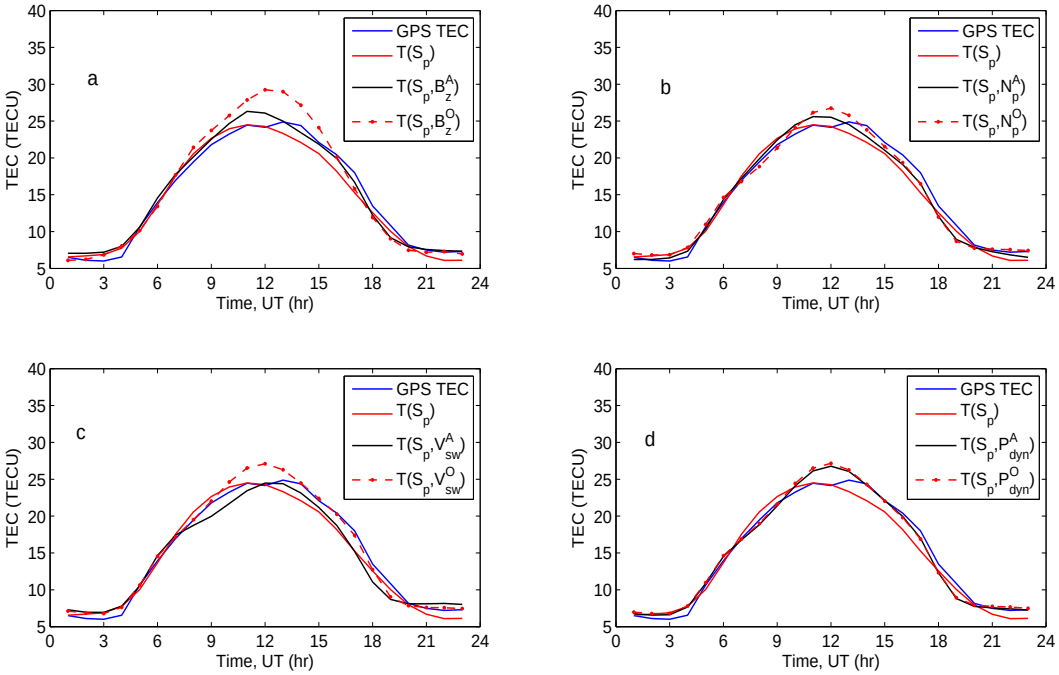


Figure 5.6: Similar to fig. 5.5, for day 285.

Table 5.2: RMSE (in TECU) between GPS TEC and predicted TEC values for days 58, 104, 185 and 285 in 2005 before and after the inclusion of the solar wind parameters from both ACE and OMNI 2 datasets.

Day	RMSE	RMSE: ACE dataset				RMSE: OMNI 2 dataset			
(in 2005)	S_p	S_p, B_z	S_p, N_p	S_p, V_{sw}	S_p, P_{dyn}	S_p, B_z	S_p, N_p	S_p, V_{sw}	S_p, P_{dyn}
58	**1.6171**	1.5873	1.3105	1.3112	1.8187	2.5423	2.5540	2.6456	2.3012
104	**2.3562**	2.7691	2.7958	3.1242	3.0064	3.3563	2.8890	2.9464	3.0383
185	**0.9472**	0.9689	1.2377	1.2506	1.1213	0.9923	1.1118	1.2943	1.1493
285	**1.2015**	1.0100	0.9443	1.2727	0.9702	2.0371	1.0824	1.0365	1.0880

Table 5.2 shows the computed RMSE values between GPS TEC and the corresponding predicted values. This table indicates that for the ACE solar wind parameters considered, B_z, N_p and V_{sw} slightly improved the prediction accuracy for day 58 while a worsening is observed for the P_{dyn} parameter. The inclusion of B_z, N_p and P_{dyn} gives improvements of $\sim 15.94\%$, $\sim 21.41\%$ and $\sim 19.25\%$ respectively for day 285 while V_{sw} slightly degrades the prediction accuracy by $\sim 5.93\%$ with respect to S_p for the ACE satellite dataset. For the OMNI 2 dataset solar wind parameters, no improvement is observed for days 58, 104 and 185 while N_p, V_{sw} and P_{dyn} slightly improves the prediction accuracy for day 285. The observed prediction accuracies in terms of both degradation and improvement are close to the values of reference when solar wind parameters are included for these quiet days. In all cases, the TEC diurnal shape is correct at almost all times of the day. Once again, these results were obtained by verifying the ENN for a GPS receiver site of which the data was not involved in model development. This confirms the previous study by Habarulema et al. (2009a) major finding of which established that given physical parameters that influence TEC variability, NNs are capable of generalising predictions both in time and space. However, there are certain latitudinal boundaries within which the NN provides accurate predictions: the further apart the stations are, the less accurate the predictions become.

5.7.1.2 During magnetically disturbed conditions

Figs. 5.7 and 5.8 show the variations of the ACE solar wind parameters and GPS TEC along with the predicted values for 06-10 and 14-18 May 2005. These particular dates were chosen because they contain days when magnetic storms occurred. These two figures are meant to specifically show the performance of the ENN model during pre-storm, storm and post-storm conditions. The prediction accuracies for these dates, where different parameters are included in ENN modelling, are shown in fig. 5.9. While the solar wind parameters show identical behaviours during the storm conditions in terms of fluctations, TEC dynamics display both positive and negative storm effects, as evidenced on days 8 and 15 May 2005.

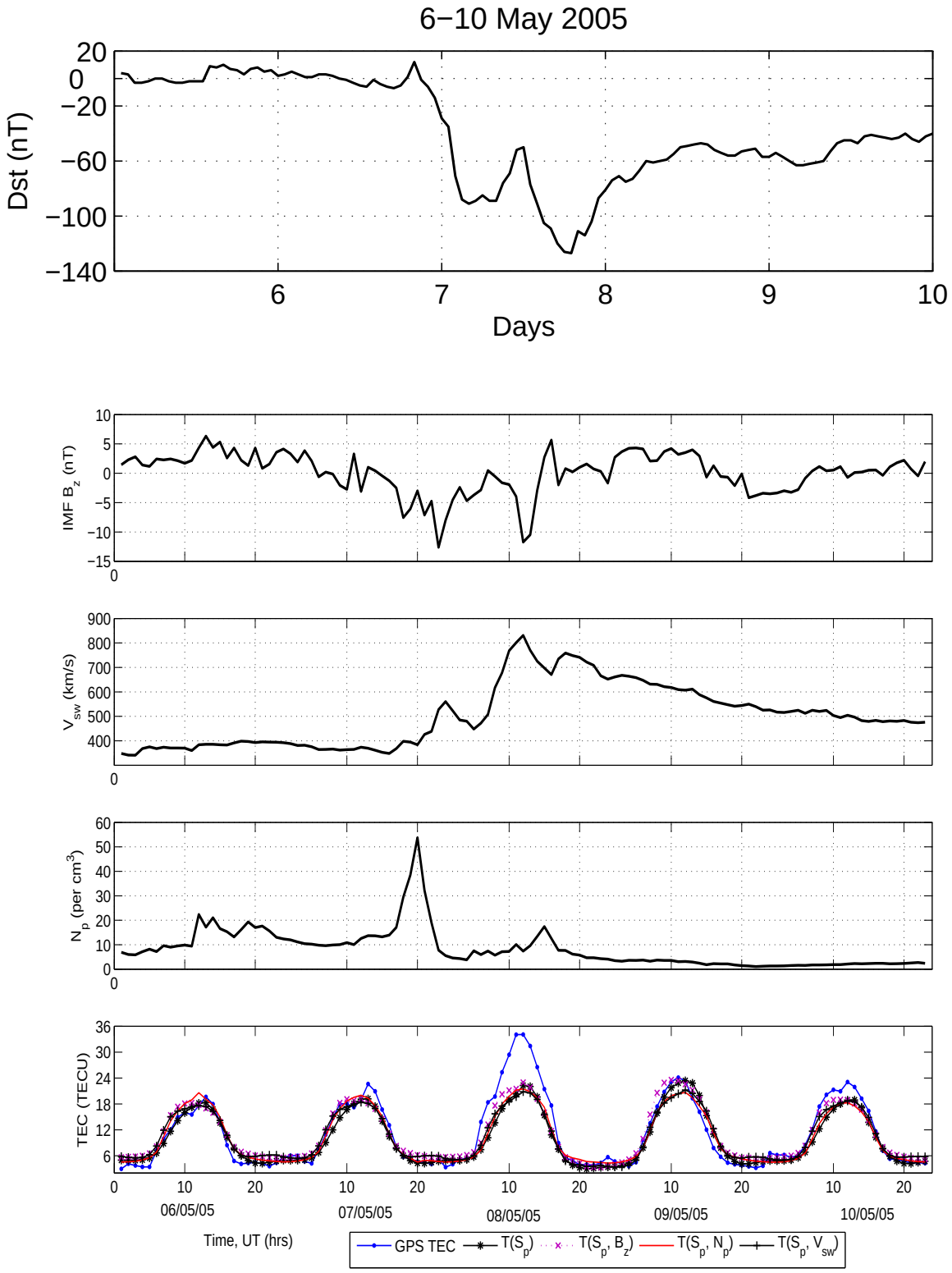


Figure 5.7: Variations of Dst and the ACE solar wind parameters and TEC values (derived and predicted) during 06-10 May 2005.

These are among the difficulties experienced in developing algorithms/models to forecast and predict most ionospheric parameters, as it is a complicated task to come up with physically

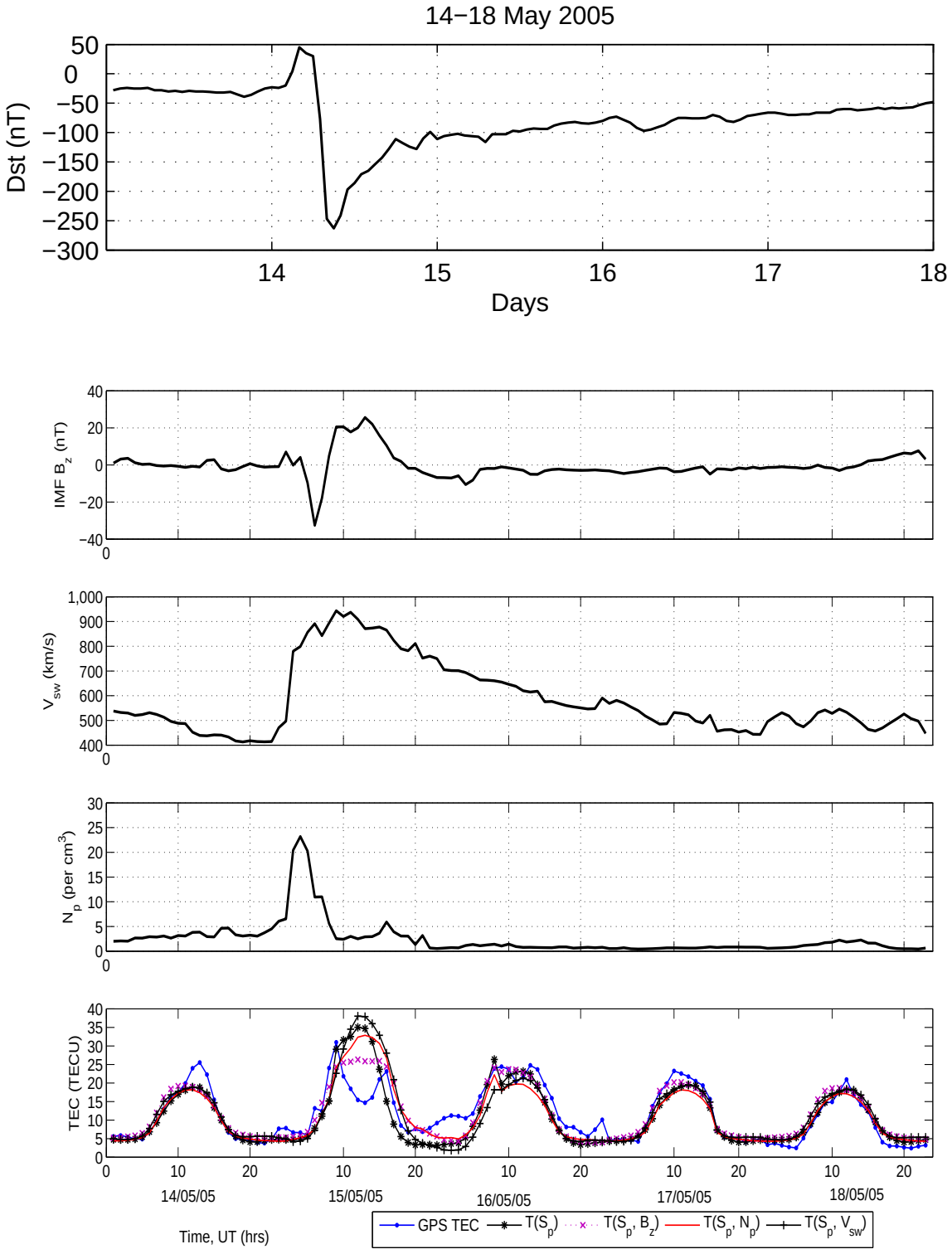


Figure 5.8: Variations of the ACE solar wind parameters and TEC values (derived and predicted) during 14-18 May 2005.

correct mathematical techniques to capture these irregular variations (e.g. Joseyln, 1995; Tulunay et al., 2006). One way to improve TEC prediction during storm conditions is to develop

a separate empirical model based on storm data and to later incorporate it in the main TEC prediction model, as was done with the IRI model. The validation of the IRI with the STORM option indicated an improvement in ionospheric parameter predictions (Araujo-Pradere and Fuller-Rowell, 2002; Araujo-Pradere et al., 2004) compared to the version that did not include the provision for disturbed conditions. This option is however not feasible for the current study due to the fact that the storm period data for the considered period may not be enough to effectively train a NN for optimum storm behaviour generalisation. Magnetic storms tend to occur more frequently during high solar activity periods compared to low solar activity periods. The latter has more data within the database used in this study. Fig. 5.7 shows that there was

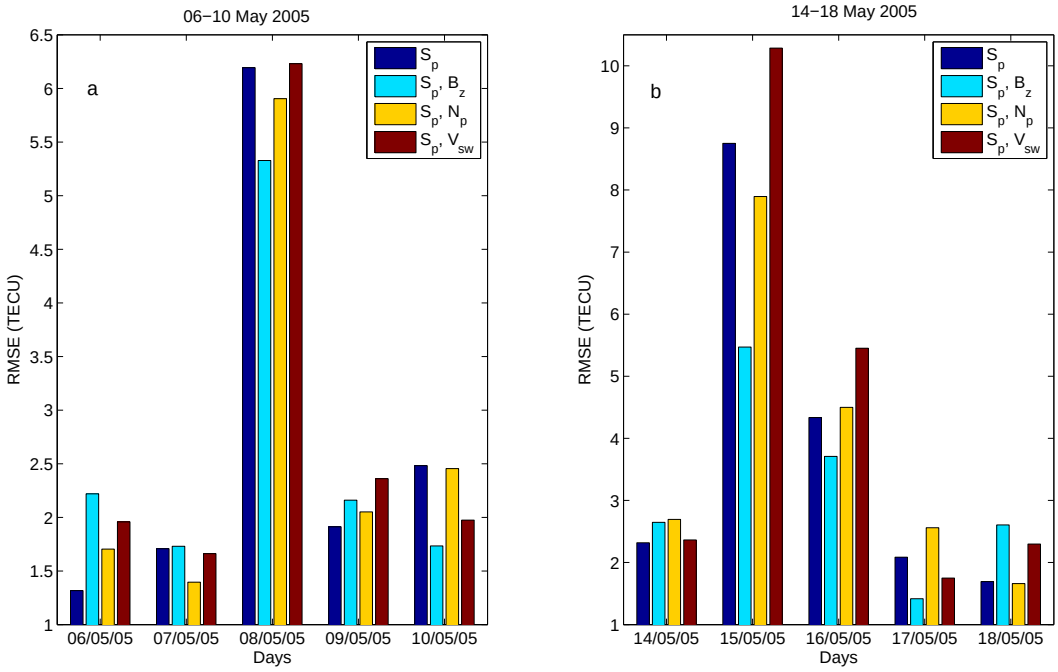


Figure 5.9: Computed RMSE values between GPS TEC and the predicted TEC as a function of all combined parameters considered for 06-10 and 14-18 May 2005.

a magnetic storm on 8 May which appears to have led to positive storm effects when the TEC values are compared to relatively quiet days before and after the storm. There was a stronger magnetic storm on 15 May 2005 (fig. 5.8) that led to the significant fluctuation of IMF B_z and sudden jump in both V_{sw} and N_p. As a result, significant TEC fluctuations are observed. For both of these storm days fig. 5.9 shows a statistical improvement after the inclusion of B_z and N_p. While the ENN correctly identifies the TEC variational pattern on 8 May, the result is different on 15 May for all parameters considered. The inclusion of P_{dyn} from both ACE and OMNI 2 datasets, as illustrated in fig. 5.10, also shows that this particular parameter doesn't lead to significant improvements in modelling storm time TEC behaviour. In addition, all combinations of different parameters strongly underpredict the magnitude of TEC values

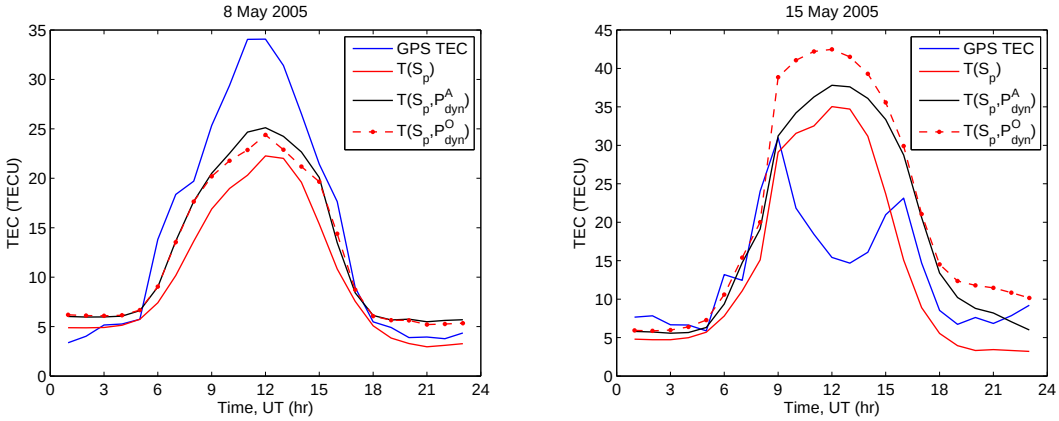


Figure 5.10: GPS TEC and corresponding predictions after the inclusion of P_{dyn} from both ACE and OMNI 2 datasets for storm days 8 and 15 May 2005.

on 8 May. In both cases, the inclusion of V_{sw} degrades the prediction accuracy, an indication that it may not be an efficient representation of solar wind in TEC modelling during magnetic storms. With respect to storm occurrences, it has been observed that V_{sw} plays a minor role in the creation of storms (Tsurutani and Gonzalez, 1997). These results clearly demonstrate that it is still difficult to accurately predict diurnal TEC behaviour during magnetically disturbed conditions based on single-station modelling.

5.7.2 Seasonal TEC predictions

Fig. 5.11 shows the comparison of GPS TEC and the predicted values over Cape Town (33.95°S, 18.47°E) at 12h00 UT in 2005. Due to a lack of data, no values are shown for November and December, 2005.

Table 5.3 shows the computed RMSE values between GPS TEC and $T(S_p)$, $T(S_p, B_z)$, $T(S_p, N_p)$, $T(S_p, V_{sw})$ and $T(S_p, P_{dyn})$ for solar wind data from both the ACE and OMNI 2 datasets. Statistically, this table reveals that all four solar wind parameters included from both datasets led to slight average improvements in TEC prediction for January, February, June and September. A remarkable observation from table 5.3 is the closeness of the prediction results (in terms of both improvements and degradation) as quantified through the RMSE method, for both ACE and OMNI 2 solar wind parameters with respect to S_p.

Fig. 5.12(a) shows the average diurnal TEC variation for January, March, June and September, 2005 representing summer, autumn, winter and spring respectively, where solar wind parameters from the ACE satellite dataset are used as inputs. The distinct observation from this figure is that all parameters underpredicted GPS TEC for almost the whole month of January, except at 03h00 UT where $T(S_p)$, $T(S_p, N_p)$, $T(S_p, B_z)$ and $T(S_p, P_{dyn})$ are greater than GPS TEC

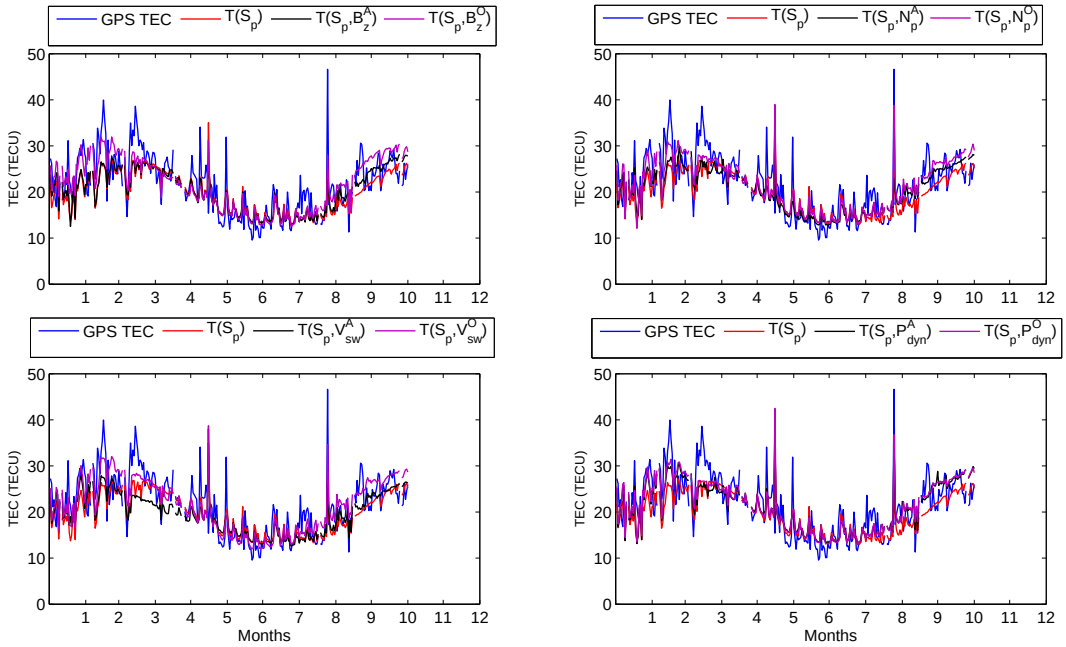
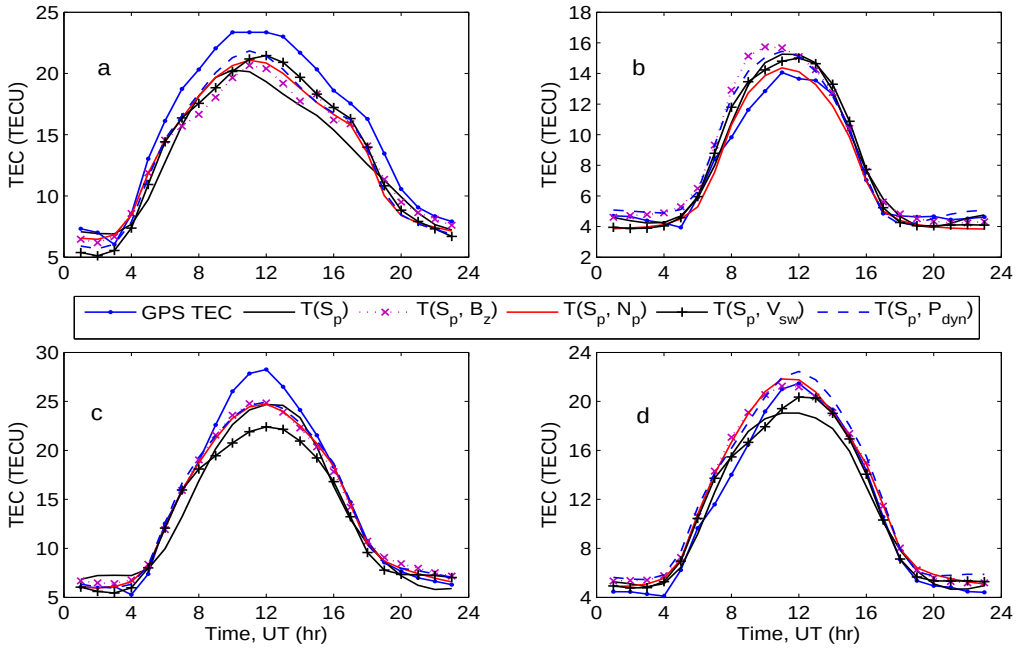


Figure 5.11: Comparison of GPS TEC and corresponding predictions at 12h00 UT in 2005.

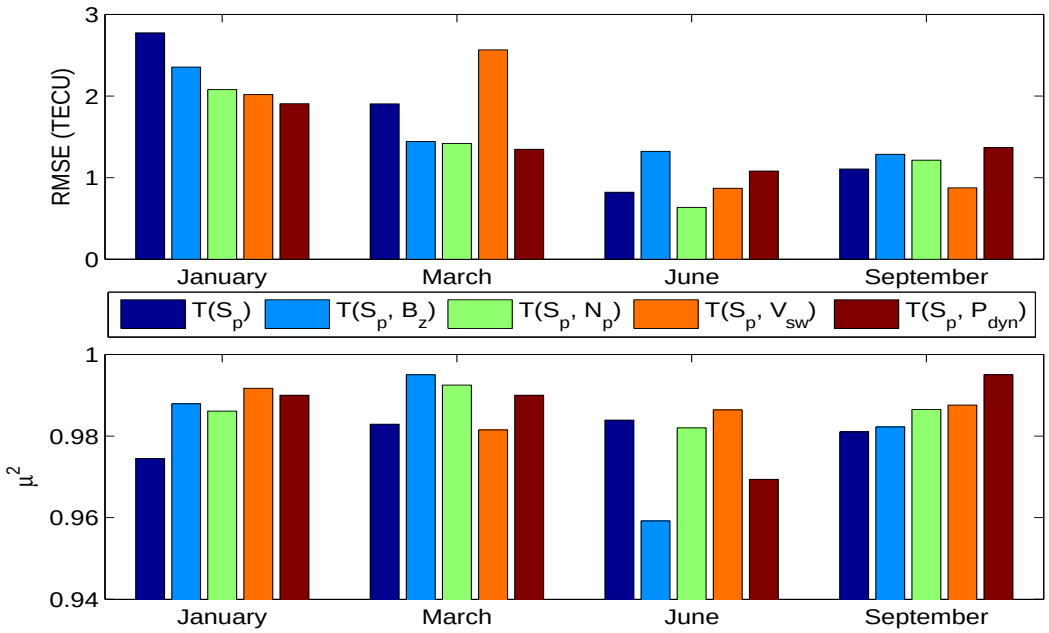
Table 5.3: RMSE (in TECU) between GPS TEC and predicted TEC values for the available months in 2005 before and after the inclusion of the solar wind parameters from both ACE and OMNI 2 datasets at 12h00 UT.

Month	RMSE	RMSE: ACE dataset				RMSE: OMNI 2 dataset			
(in 2005)	S_p	S_p, B_z	S_p, N_p	S_p, V_{sw}	S_p, P_{dyn}	S_p, B_z	S_p, N_p	S_p, V_{sw}	S_p, P_{dyn}
1	**5.5471**	5.0578	4.3524	4.1936	4.3254	3.7369	4.4346	3.8946	4.1686
2	**6.5186**	6.3734	5.8721	5.7605	5.2727	5.2265	5.2656	5.3205	5.1747
3	**5.7345**	5.5850	5.6514	7.3874	5.7194	5.0707	4.7302	4.6108	5.1790
4	**2.0816**	2.2099	2.3469	3.6802	2.3936	2.2184	2.5398	2.5040	2.4549
5	**5.5885**	4.7351	5.7753	6.2910	5.7167	5.3590	6.0113	6.1183	6.3422
6	**2.4282**	2.3319	1.9745	2.2795	2.3460	2.1372	2.2322	1.8420	2.2613
7	**2.2290**	2.2684	2.4804	2.5171	2.3007	2.3366	2.2820	2.4266	2.2973
8	**5.4885**	5.5077	5.4013	5.0770	3.9443	4.7844	3.5383	3.9382	3.7338
9	**4.2888**	3.6763	3.7207	3.7255	3.8793	3.8880	3.9483	4.0590	3.7684
10	**2.5702**	2.7704	2.8691	2.5784	3.4082	4.2038	3.6132	3.4314	3.5426

by $\sim$ 0.871 TECU, 0.779 TECU, 0.629 TECU and 0.013 TECU respectively, and 04h00 UT where T(S_p, N_p) $\simeq$ T(S_p, B_z) $\simeq$ GPS TEC. There is an underprediction for March between $\sim$ 08h00 UT and 16h00 UT. Fig. 5.12(b) represents the RMSE between derived GPS TEC and its predicted values as a function of all considered parameters. Represented in fig. 5.12(b) are also the coefficients of determination (square of correlation coefficients) between GPS TEC and TEC values as predicted by different parameter combinations. Fig. 5.13(a) shows the TEC

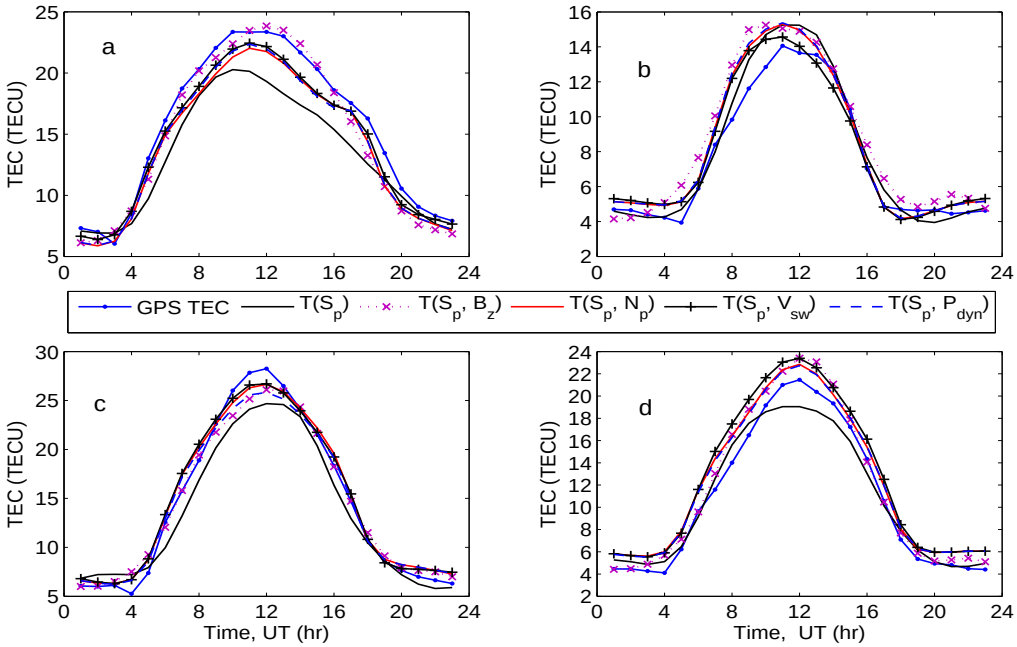


(a) Monthly average GPS TEC for each hour and the corresponding predictions with all parameters used for (a) summer (January), (b) winter (June), (c) autumn (March) and (d) spring (September) in 2005.

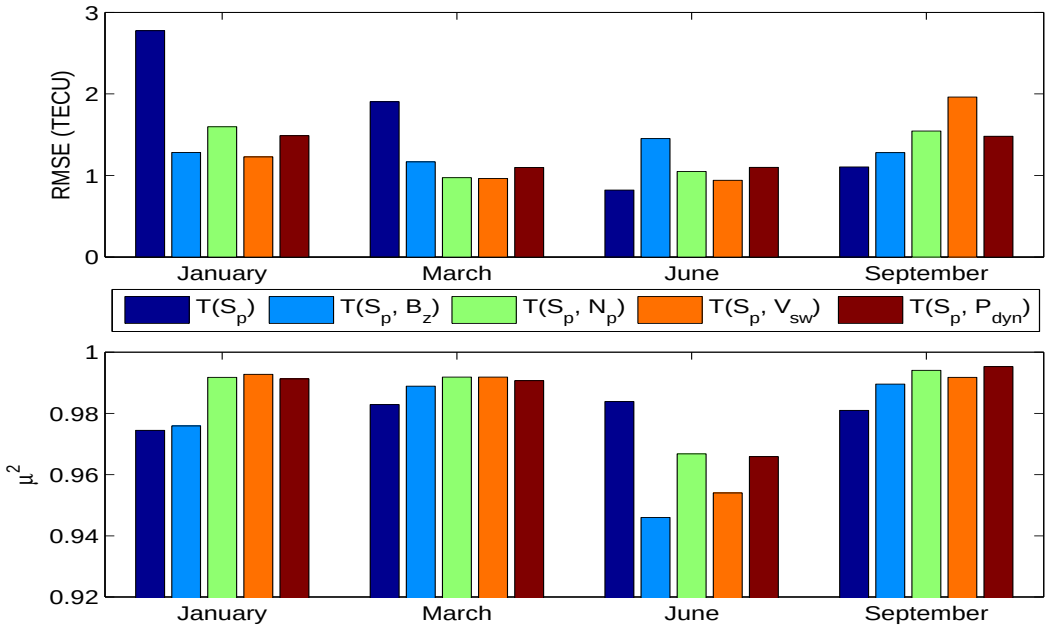


(b) RMSE and μ^2 values for the months representing autumn, winter, spring and summer in 2005.

Figure 5.12: Mean monthly hourly GPS TEC and corresponding predictions with the statistical comparison of the prediction accuracies in 2005. ACE solar wind parameters were used.



(a) Monthly average GPS TEC for each hour and the corresponding predictions with all parameters used for (a) summer (January), (b) winter (June), (c) autumn (March) and (d) spring (September) in 2005.



(b) RMSE and μ^2 values for the months representing autumn, winter, spring and summer in 2005.

Figure 5.13: Similar to fig. 5.12, but with OMNI 2 dataset solar wind data.

variations (both GPS and corresponding predictions) when the OMNI 2 dataset solar wind parameters are used in modelling. Fig. 5.13(b) shows the statistical analysis of the prediction

accuracy for each solar wind parameter for these months representing summer, winter, autumn and spring. A comparison between the performances of these parameters between the ACE and OMNI 2 datasets is shown in table 5.4, in terms of relative RMSE deviations computed as

$$R_D = \frac{R^{S_p} - R^{swp}}{R^{S_p}} \tag{5.8}$$

R_D: relative RMSE deviation,
R^{S_p}: computed RMSE between GPS TEC and TEC predicted as a function of S_p,
R^{swp}: RMSE between GPS TEC and predicted TEC as a function of S_p and each of the solar wind parameters

A negative RMSE deviation indicates a degradation in prediction after the inclusion of the solar wind parameter. This table indicates that almost all solar wind parameters improved

Table 5.4: Comparison of the solar wind parameters' performance from both ACE and OMNI 2 datasets for months representing seasons.

Month (in 2005)	R_D: ACE dataset				R_D: OMNI 2 dataset			
	B_z	N_p	V_{sw}	P_{dyn}	B_z	N_p	V_{sw}	P_{dyn}
January	0.1516	0.2506	0.2728	0.3131	0.5382	0.4244	0.5575	0.4639
March	0.2427	0.2550	-0.3468	0.2927	0.3873	0.4889	0.4948	0.4236
June	-0.6119	0.2256	-0.595	-0.3161	-0.77	-0.2783	-0.1464	-0.3392
September	-0.1631	-0.988	0.2090	-0.2385	-0.1578	-0.3975	-0.7749	-0.34

TEC predictions for January and March, with the OMNI 2 dataset providing more accuracy. A degredation is only observed for the ACE V_{sw} inclusion for March. In June and September, only N_p and V_{sw} respectively showed an improvement in TEC predictions for the ACE satellite dataset, while the OMNI 2 dataset revealed a degredation for all the solar wind parameters. Girish et al. (1997) obtained a significant correlation between the variations of solar wind parameters and TEC during the sunspot maximum in winter and spring for the Northern hemisphere. The analysis done in this study also indicates a statistical correllation between TEC variations and at least one of the solar wind parameters investigated in all seasons by means of ENN modelling. These results indicate a relative agreement for N_p and V_{sw} during the declining phase of the solar cycle in the Southern hemisphere, which may be a confirmation that the solar wind dynamics influences TEC variation at both solar activity levels and in both hemispheres.

5.8 Summary

This chapter has attempted to explain the quantification of solar wind effects on GPS TEC using ENN modelling. GPS TEC was predicted as a function of $D_n, H_r, R4$ and $A8$, and the results used as a benchmark to try and quantify solar wind influence on GPS TEC. Statistical analyses revealed both improvement and degradation in TEC prediction accuracy for different parameter combinations and for different days and seasons. An analysis was performed for both quiet and storm conditions. Although the inclusion of B_z and N_p improved the prediction accuracy for both stormy days (8 and 15 May), an underprediction was observed for 8 May, while the variational pattern was incorrect for 15 May. These results show that it is still difficult to predict TEC variability during magnetic storms. For both diurnal and seasonal TEC-modelled results, the degradation or improvement observed was near the values predicted by the standard parameters and this study has lent more weight to the belief that diurnal, seasonal, solar and magnetic variability may be the major determinants of TEC variability. As a result, the next chapter doesn't take into account the solar wind contribution to TEC variability in regional TEC modelling.

Chapter 6

Southern African TEC Prediction (SATECP) Model

This chapter describes the final version of the GPS-based TEC prediction model that was developed as a function of magnetic and solar activities, diurnal and seasonal variations and the geographical position of the GPS receivers. It represents a combination of all the feasibility studies carried out in the quest for a TEC prediction model for Southern Africa.

6.1 GPS infrastructure used

Within South Africa continuous GPS data recording started in 2000, and the availability of data since that year is station dependent. At the time of writing this thesis, South Africa had a network of approximately 50 GPS receiver stations. Data from 14 stations were used in model development, with other 8 stations providing data for the final verification of the NN model. Fig. 6.1 shows the location of the stations used in the training, validation/testing and verification processes. The choice of the stations was based on two factors (Habarulema et al., 2009b), namely;

1. The effort to cover the entire South Africa: This factor is supported by the study in chapter 4 which clearly indicates the validity of using single-station NN models to model/predict TEC variability at stations where data was not used in training, within a latitudinal difference range of $\sim$ 1 - 3 degrees.

2. Data availability and consistency: Due to infrastructure and hardware problems, there are data gaps within the data. Therefore, it was necessary to identify stations with a reliable data history and minimal data gaps for this study.

Other underlying factors considered in data selection included the manifestation of TEC variability according to expected trends such as values being high or low during high/low solar

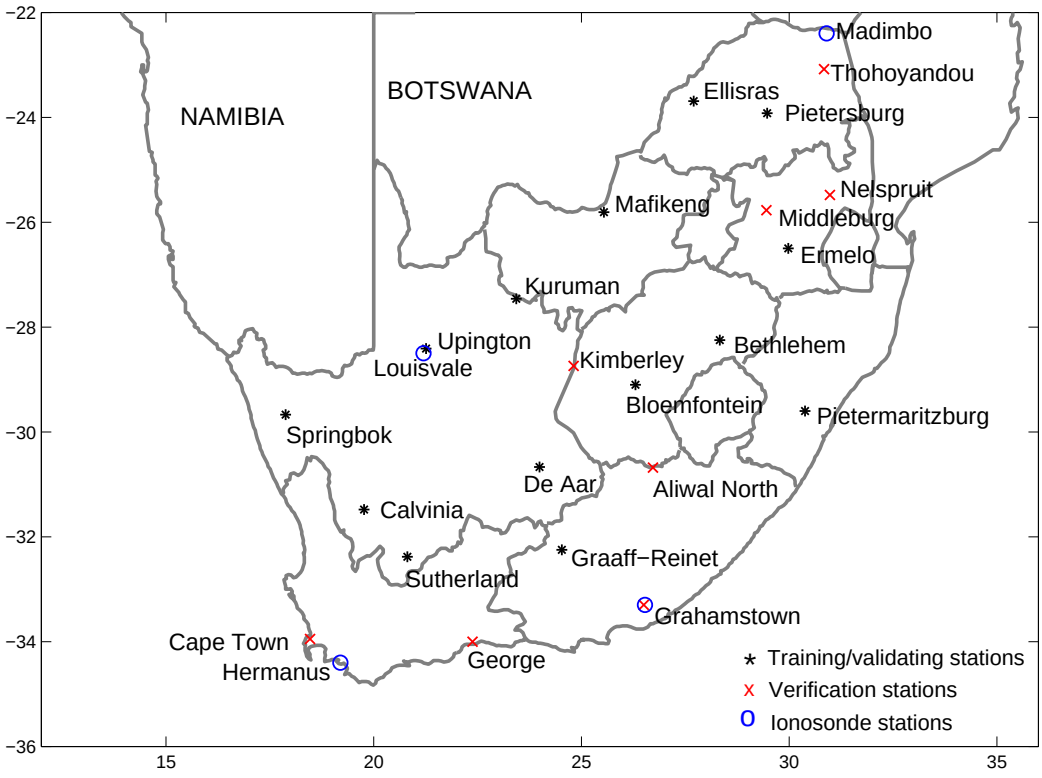


Figure 6.1: Map of South Africa showing training and verification stations. The locations of the ionosonde stations are also shown.

activity periods, and the manifestation of viable diurnal and seasonal variations. For example, daily TEC values reach maximum at $\sim$ 02h00 local time, and seasonally, general trends indicate that TEC is higher during summer than in winter. Any processed data displaying negative TEC values were removed and flagged as data errors since negative TEC values are not physically reasonable.

6.2 Data range and consideration

GPS TEC data at a temporal resolution of 1 minute for an interval of at most 7 years (2000-2006) was derived using the ASHA algorithm. GPS stations have different available data ranges and tables 6.1 - 6.2 show the geographical positions of GPS receiver stations used in the model's development and validation (training, testing/validation and verification processes). It is difficult to find a single year with all data for all times of the day and in some cases data gaps exist for several days. One of the reasons for carrying out this study was to fill the data gaps with predictions based on actual data from the same regions. The data range which was opted

Table 6.1: Geographical locations of the GPS receiver stations within South Africa and the time periods for which training data was available for use in the development of the SATECP model.

GPS receiver station	Station code	Latitude (° S)	Longitude (° E)	Years for which data was available
Bethlehem	BETH	28.25	28.33	2001-2006
Bloemfontein	BFTN	29.10	26.30	2002-2006
Calvinia	CALV	31.48	19.77	2002-2006
De Aar	DEAR	30.67	23.99	2002-2006
Ellisras	ERAS	23.69	27.70	2001-2006
Ermelo	EMLO	26.50	29.98	2002-2006
Graaff-Reinet	GRNT	32.25	24.53	2003-2006
Kuruman	KMAN	27.46	23.43	2002-2006
Mafikeng	MFKG	25.81	25.54	2002-2006
Pietermaritzburg	PMBG	29.60	30.38	2001-2006
Pietersburg	PTBG	23.92	29.47	2002-2006
Springbok	SBOK	29.67	17.88	2002-2006
Sutherland	SUTH	32.38	20.81	2000-2006
Upington	UPTN	28.41	21.26	2001-2006

Table 6.2: Geographical location of the GPS receiver stations within South Africa and the year for which the verification of the SATECP model was carried out.

GPS receiver station	Station code	Latitude (° S)	Longitude (° E)	Testing data
Aliwal North	ANTH	30.68	26.72	2003
Cape Town	CPTN	33.95	18.47	2002
George	GEOR	34.0	22.38	2002
Grahamstown	GRTN	33.3	26.5	2005
Kimberley	KLEY	28.74	24.81	2002
Middleburg	MBRG	25.77	29.45	2003
Nelspruit	NSPT	25.48	30.98	2002
Thohoyandou	TDOU	23.08	30.84	2005
Sutherland	SUTH	32.38	20.81	2007

for in the model development includes a high solar activity period (e.g. years 2000 and 2001) and periods which fall in the declining phase of the solar cycle (e.g. years 2005 and 2006). Thus the application of the model for data recovery based on predictions can be considered valid for these interval ranges.

6.3 Model development

Recent studies described in Habarulema et al. (2007, 2009a,b) show the capabilities of NNs for TEC predictions within South Africa by using multiple inputs through an input-output mapping

process. Results reported in Habarulema et al. (2007) and Habarulema et al. (2009a) dealt with single-station studies for hourly TEC modelling/predictions, while Habarulema et al. (2009b) presented first results from a national GPS-based TEC prediction model for South Africa for the period 2000 - 2004 comprising 1-minute data from 10 GPS receiver stations. Single-station modelling implied that the geographical dependance of the TEC parameter was not considered, as described in chapters 4 and 5. The new version of the model (SATECP) described in this chapter utilises data from a period of 7 years (2000 - 2006) and extends the number of GPS receiver stations considered to 14 for more spatial and temporal coverage.

Factors which influence TEC variability were categorised into periodic, random and positional components. The periodic component (P_c) corresponds to the diurnal and seasonal variations exhibited in TEC behaviour, i.e $P_c == H_r, D_n$ (e.g. Stankov et al., 2001) and acts as the initial modelling parameters for the TEC time series simulation. The random component (R_c) represents the complex and strong changes that influence TEC variability as a result of solar and geomagnetic activities (Stankov et al., 2001). The positional component (G_p) corresponds to the geographical positions of the GPS receiver stations included and is represented by geographic latitude and longitude (G_p ==lat,long). The predicted outcome is vertical TEC as a function of 8 parameters (inputs) represented by their respective variables as shown in Table 6.3.

Table 6.3: Different factors used during model development and their numerical representations.

	Factor	Its representation
1.	Seasonal variation	Day number, D_n (1 - 365 or 366 for leap years)
2.	Diurnal variation	Hour, H_r (1-24, minutes were converted to hrs)
3.	Solar activity	Sunspot number
4.	Magnetic activity	Magnetic index values from a local observatory, Hermanus
5.	GPS receivers' location	Geographical latitude and longitude (degrees)

The final predicted TEC can be mathematically expressed as

$$\left.\begin{array}{l} T_p \equiv f(P_c, R_c, G_p) \\ \text{where} \\ P_c == \{D_n, H_r\} \\ R_c == \{R4, A8\} \\ G_p = \{lat, long\} \end{array}\right\} \quad (6.1)$$

T_p: TEC predicted by the SATECP model

lat, *long*: latitude and longitude representing the geographical position (G_p) of the GPS receivers used in the model development.

For the temporal input, minutes are converted to hours for consistency and later split into cyclical trigonometrical components in order to take into account the adjacency of the last hour of the day and the first hour of the next day (e.g. Williscroft and Poole, 1996; Poole and Poole, 2002; Senalp et al., 2008). For the same reason, the day number of the year (representing seasonal variation) is split into its two trigonometrical components. As determined in Habarulema (2007), $R4$ is the average of the preceeding 4 months' daily sunspot number values while $A8$ is the average of the previous 1-day magnetic index values derived from Hermanus magnetic observatory (34.43°S,19.23°E). During training, the NN software uses $R4, A8, D_n, H_r$, and G_p to establish the variational relationship between TEC and solar activity, magnetic activity, seasonal variation, diurnal variation and the geographical position of the GPS receivers respectively, using non-linear mathematical modelling and interconnectivity between input and output parameter mapping with the aid of different nodes/units within the respective layers. To illustrate how different input parameters influence the simulation of TEC time series behaviour, six separate experiments were carried out, using the three categorised components. Keeping the periodic component constant, the random and positional component representations were included in the NN and modelling was performed to determine the level of influence of each input parameter on storm and quiet-time TEC fluctuations. Fig. 6.2 shows the computed RMSE values between the modelled and actual TEC values for the considered storm and quiet periods over selected stations within the region of interest. Generally, results show that the greatest contributor to the accuracy of TEC modelling/predictions is the solar activity parameter. This observation holds mainly for relatively quiet conditions (e.g. fig. 6.2d). For the G_p component, the longitudinal and latitudinal contributions are close to each other, although in the final model both were included to maximise spatial contribution. Removal of the geomagnetic activity representation degrades the accuracy in the modelled TEC especially on storm days (e.g. 2 October 2002, fig. 6.2c) over KLEY and SUTH with the exception of 29 October 2003 as shown in fig. 6.2b.

6.3.1 Final NN architecture

Generally, there is no standard for defining the number of hidden units/nodes within the network architecture. Users process and structure data in different ways for optimisation and the general method for determining the number of hidden nodes can best be described as "trial and error". Several NNs were trained and verified in an effort to find a suitable architecture for the optimum solution. Due to the amount of data under consideration, the training time is longer for single hidden layered networks and becomes shorter for multilayered feed forward networks. To find the optimum solution, all inputs are fed into the network and hidden layers varied with varying hidden nodes, generating predictions using data that was not included in the network and finally performing a statistical analysis to determine the suitable architecture.

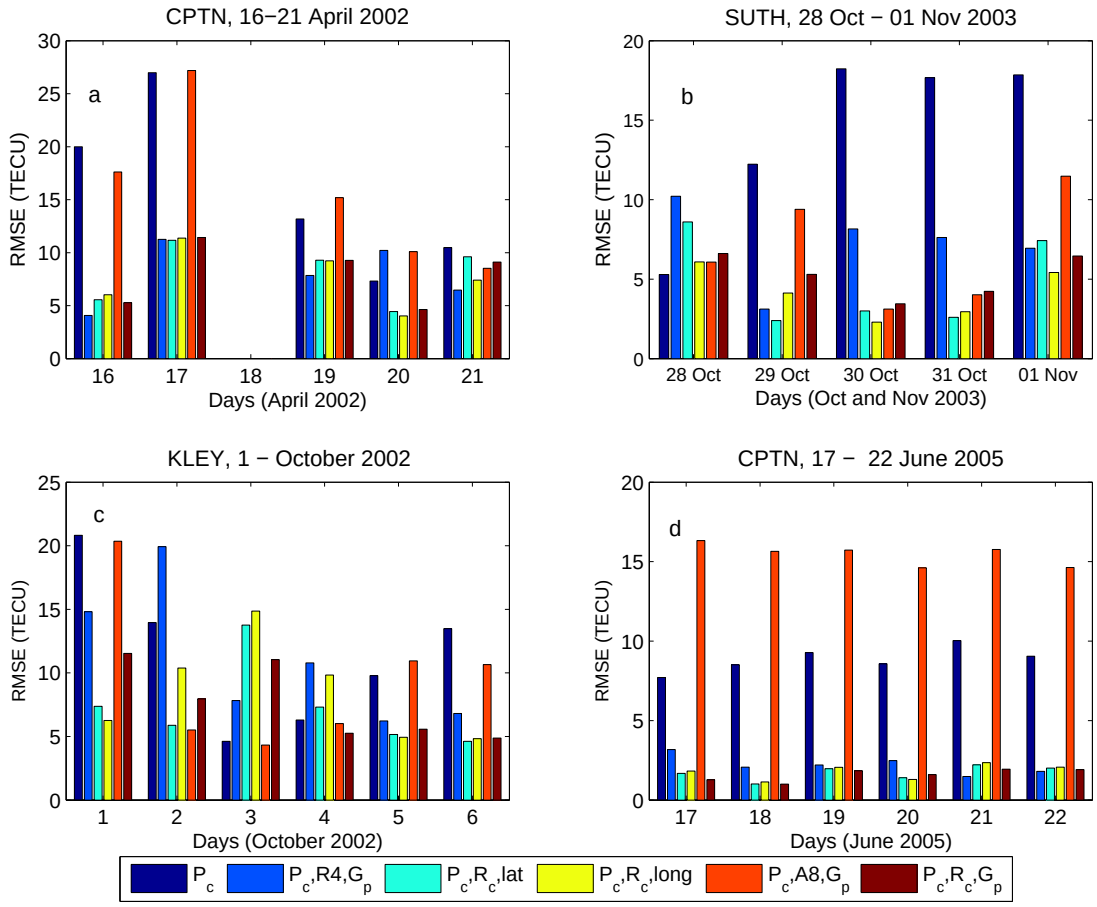


Figure 6.2: An illustration of the influence of NN input parameters on TEC time series behaviour.

During training, memorisation is avoided by using the testing/validation data pattern within the training process and halting training when the mean square error of the validation data pattern starts increasing. The testing/validation data pattern is derived from the combination of the overall data from all the stations considered as 30% of the entire dataset after its random sorting.

With 8 inputs and about 17 million data points (precisely 16877752), the suitable architecture consisted of one input layer ($N_i = 8$), two hidden layers ($N_h = 11$ in each layer) and one output layer ($N_o = 1$), where N_i, N_h and N_o are the nodes in the input, hidden and output layers respectively. More hidden layers enable the network to extract statistical terms of high order, especially when the input space is quite large (Yilmaz et al., 2009). In this study, each 1-minute data point or set consists of 8 inputs and one target value (TEC). With a NN configuration of 8:11:11:1, the approximated optimum solution was obtained using the standard back propagation algorithm and implementing the feed forward multi-layer perceptron (MLP) network in a topological order, employing 250 iterations. Normalisation was done by transforming data for all input parameters and corresponding TEC values into the range of

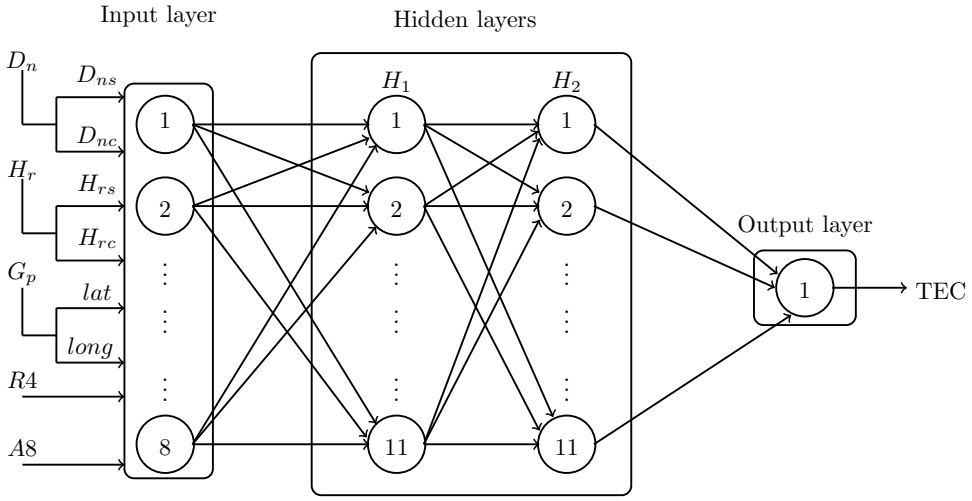


Figure 6.3: A schematic illustration of the final NN architecture.

(-1,1) for the purpose of minimising the training and testing time of the network. Fig. 6.3 shows the full picture of the interconnectedness of the neurons and layers within the network that was used to generate TEC predictions in this chapter.

6.4 Diurnal verification of SATECP model

The inputs are 1-minute interval data parameters, thus the lower limit temporal resolution at which predictions can be generated. However flexibility is allowed where modelled/predicted TEC data can assume values more than 1 minute. In assessing the diurnal prediction capabilities of the SATECP model, TEC variabilities over the verification stations (table 6.2) were generated for equinox and solstice days. In cases where data for these particular days were not easily available, the closest days with data were used (Habarulema et al., 2009b). Predictions were at times generated for days before or after the equinox/solstice day to compare values of different verification GPS stations. Table 6.4 shows a summary of equinoix and solstice days in 2002, 2003 and 2005 intended for verification of TEC predictions. Figs. 6.4 and 6.5 show GPS TEC and predicted values over selected verification stations for days representing equinox and solstice days in 2002 and 2003 respectively. TEC values generated by the SATECP model for 2002 indicate the performance of this model near the high solar activity period, while 2003 predictions take into account the period of the declining phase of the solar cycle. Table 6.5 shows the computed RMSE values between GPS TEC and predicted TEC as obtained from the first attempted national TEC prediction model and the current version of the SATECP model over ANTH and MBRG in 2003.

Table 6.4: Dates when equinox and solstice days occured during 2002, 2003 and 2005.

Year	Equinox days	Solstice days
2002	20 March and 23 September	21 June and 22 December
2003	20 March and 23 September	21 June and 22 December
2005	20 March and 22 September	21 June and 21 December

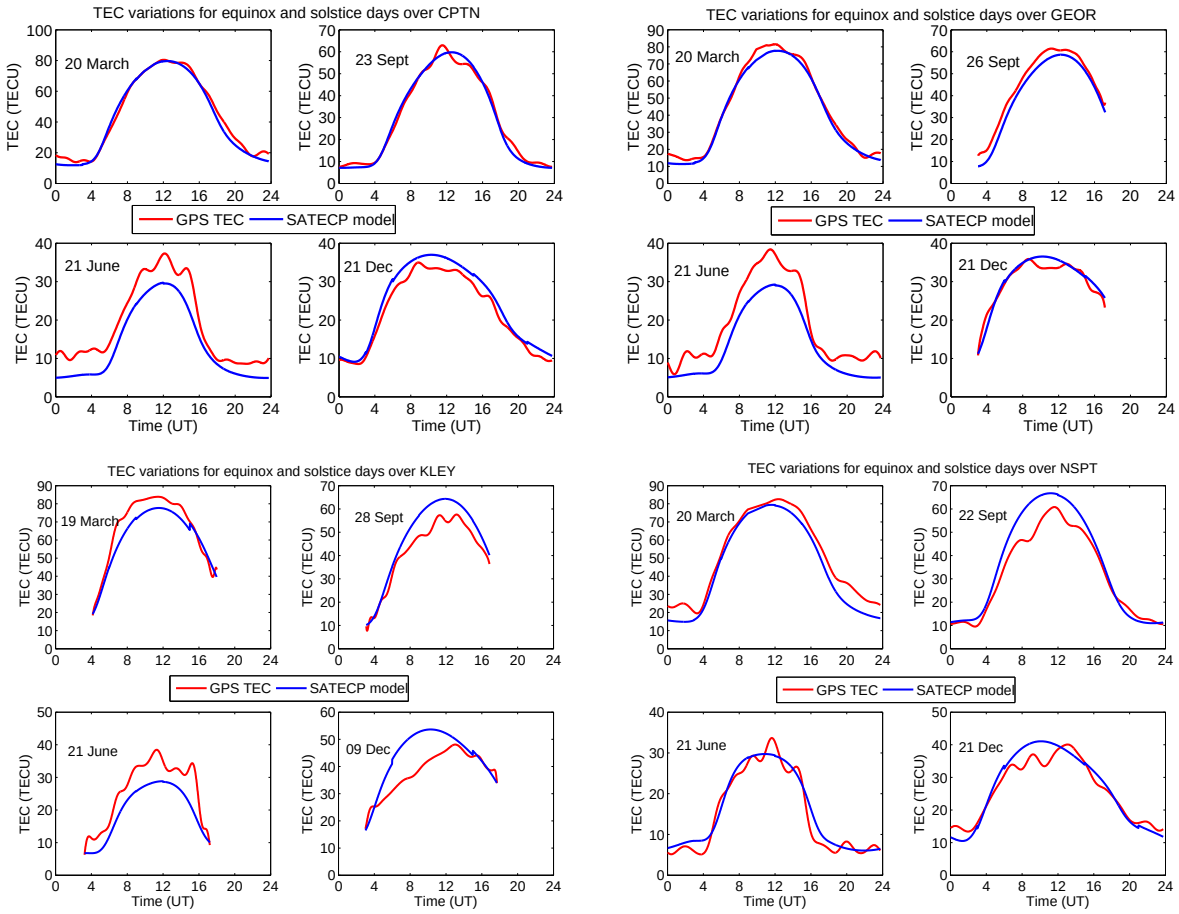


Figure 6.4: 1-minute GPS TEC with the corresponding SATECP model predictions for days representing equinox and solstice days in 2002.

As indicated in table 6.5, the SATECP model included more data than the first attempted TEC prediction model. The addition of more data leads to a general significant improvement in the prediction accuracy for most of the days representing equinoxes and solstices, particularly for June 21 and January 17 over both receiver stations, and September 23 for ANTH. In a few cases the prediction accuracy remains unchanged or deteriorates.

A fair sample analysis of the prediction results can be made by looking at TEC values for the same day. By considering this factor, the differences in RMSE values (Δ RMSE) for June 21, 2002 is $\sim$ 3.148 TECU (CPTN and NSPT) compared to $\sim$ 0.661 TECU for June 21,

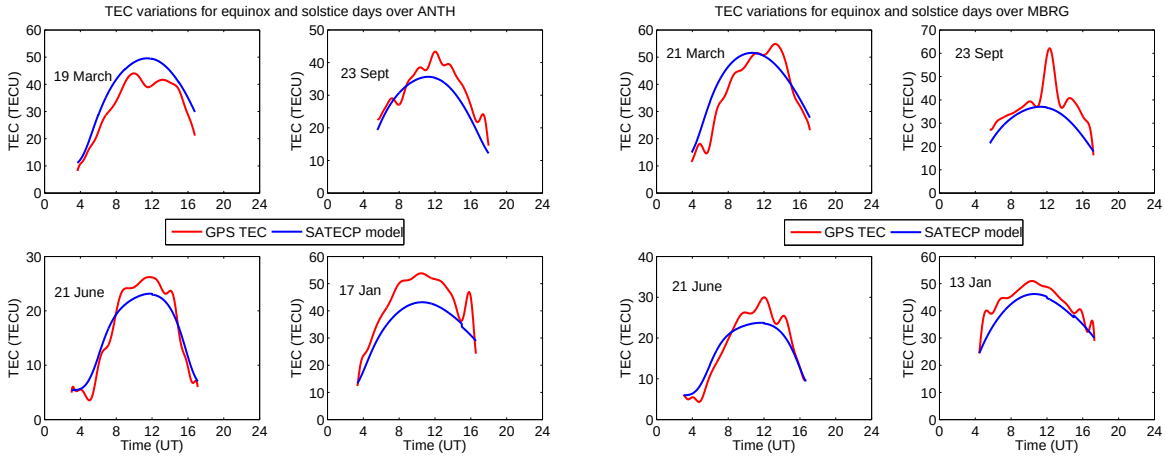


Figure 6.5: Similar to fig. 6.4 but for 2003.

Table 6.5: RMSE values between GPS TEC and predicted TEC values computed by Habarulema et al. (2009b) and the SATECP model over ANTH and MBRG.

Verification dates representing equinoxes and solstices (in 2003)	RMSE (TECU) for the 1^{st} model with data: 2000-2004		RMSE for the SATECP model data expanded to 2000-2006	
	ANTH	MBRG	ANTH	MBRG
March 19 (ANTH)/21 (MBRG)	3.790	6.048	6.203	6.645
Sept 23 (both ANTH&MBRG)	7.205	8.726	4.659	9.136
June 21 (both ANTH&MBRG)	3.873	3.883	2.342	3.003
Jan 17 (ANTH)/13 (MBRG)	10.507	5.114	8.302	4.396

2003 over verification stations ANTH and MBRG, with latitudinal differences of 8.47° and 4.91° respectively. Apart from the latitudinal aspect, this is an expected trend as it also demonstrates the different solar activity levels for the considered periods. Fig. 6.6 shows the GPS TEC variations along with the corresponding predictions by the SATECP model for days representing equinoxes and solstices in 2005 over GRTN and TDOU. These are particulary interesting verification sites for statistical comparison, since TDOU is the northernmost site and GRTN lies in the southern part of the region of interest. For this reason, the generation of prediction results on similar days for both sites is necessary. However, the unavailability of GPS data during the December led to the verification of the SATECP model with January data for the summer solstice. Only days with complete data sets (equinox days) were analysed to determine the model's latitudinal estimation accuracy.

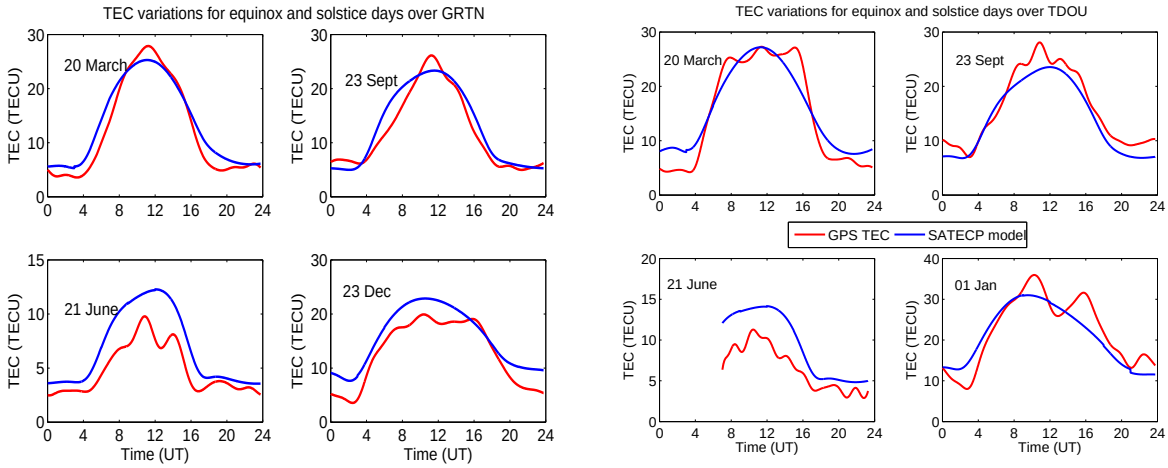


Figure 6.6: Similar to fig. 6.4 but for 2005.

The RMSE estimates between GPS TEC and predicted TEC for the autumn equinox (20 March 2005) are ~ 2.857 TECU and 1.983 TECU over TDOU and GRTN respectively. These values change to ~ 2.218 TECU and 2.089 TECU for the spring equinox (23 September 2005) and this demostrates how close and comparable the predictions are on similar days over these verification stations with a latitudinal separation of $\sim 10°$. The results and their statistical analysis in this section demonstrate that the SATECP model is capable of identifying and generalising diurnal TEC variability patterns at all the latitudes considered. Having confirmed this capability, further investigation of TEC predictability with respect to other conditions was undertaken. These will be discussed in the next sections and chapter.

6.5 SATECP model verification with other data sources

Two methods were used to perform an "independent" validation of the SATECP model. In this context, independent means using data that was collected/estimated/derived with techniques different from the ones employed in the SATECP model. Firstly, by taking TEC values estimated by the ASHA algorithm as the reference, the developed SATECP model was used to predict GPS TEC and compare the results to TEC data derived from the UNB-IMT model. In the second independent validation option, the SATECP model predictions were compared with ionosonde-derived TEC (ITEC). The electron density profiles obtained from ionosonde measurements (by inverting bottomside ionograms) can be used to derive the ITEC up to an altitude of about 1000 km. This is achieved by a combination of the measured bottomside and modelled topside ionospheric profiles obtained from the ARTIST4 scaling software used within the Digisonde systems (Reinisch and Huang, 2001; Reinisch et al., 2004). TEC values above the ionosonde locations were predicted and compared to TEC values derived from actual ionosonde measurements. Only three ionosonde stations (Grahamstown, Madimbo and Louisvale) were

Table 6.6: Geographical location of the ionosondes in South Africa.

Station name	Station code	Latitude (° S)	Longitude (° E)	Geomagnetic latitude
Grahamstown	GRTN	33.30	26.53	-34.30
Louisvale	LVLE	28.50	21.20	-28.53
Madimbo	MDBO	22.40	30.90	-24.28
Hermanus	HERM	34.43	19.23	-42.67

considered since the Hermanus ionosonde was only installed in 2008 and thus does not cover the data period of the developed SATECP model. Table 6.6 shows the geographic location of the South African ionosonde stations.

6.5.1 SATECP and UNB-IMT models compared

As mentioned in chapter 3, the UNB-IMT algorithm was adapted to characterise the solar cycle effects on GNSS-derived TEC over Southern Africa using data from IGS stations as well as the South African GPS receiver network (Moeketsi et al., 2007; Moeketsi, 2008). The validation of the UNB-IMT model over South Africa was performed using ITEC (from South African ionosonde stations) and IRI-2001 model predictions, and all these techniques were generally found to have a good degree of agreement during solar maximum (2002) and minimum (2005) conditions respectively (Moeketsi, 2008). Here a quick comparison of the two models with reference to GPS TEC derived with the ASHA algorithm is performed for receiver stations located at East London, ELDN (33.04° S, 27.83° E) and Phalaborwa, PBWA (23.95° S, 31.13° E) for 09 April and 27 August 2002 respectively. Fig. 6.7 shows the diurnal TEC comparison between TEC derived from the ASHA algorithm, UNB-IMT model and the predicted values by the SATECP model. The temporal resolutions are 1 minute for the ASHA algorithm and SATECP model, and 5 minutes for the UNB-IMT model. This comparison indicates that the use of different models or techniques allows for continuous data coverage of particular periods (as depicted over PBWA, 27 August 2002), which is useful in developing reliable ionospheric models. Over PBWA, UNB-IMT provides slightly lower TEC values than the ASHA algorithm, while by visual inspection, TEC predictions from the SATECP model appear similar to the average of the two models. There are differences in TEC magnitude between UNB-IMT and ASHA. Notable differences between these two approaches include

- The use of spherical harmonics in TEC estimations for ASHA as opposed to spatial linear approximation of stochastic parameters in a Kalman filter optimisation for the UNB-IMT.
- UNB-IMT and ASHA use 10° and 15° elevation cutoff angles respectively for minimising multipath errors.

- Use of different ionospheric thin shell heights (350 km and 400 km for ASHA and UNB-IMT respectively) in their mapping functions that transform slant TEC to vertical TEC.

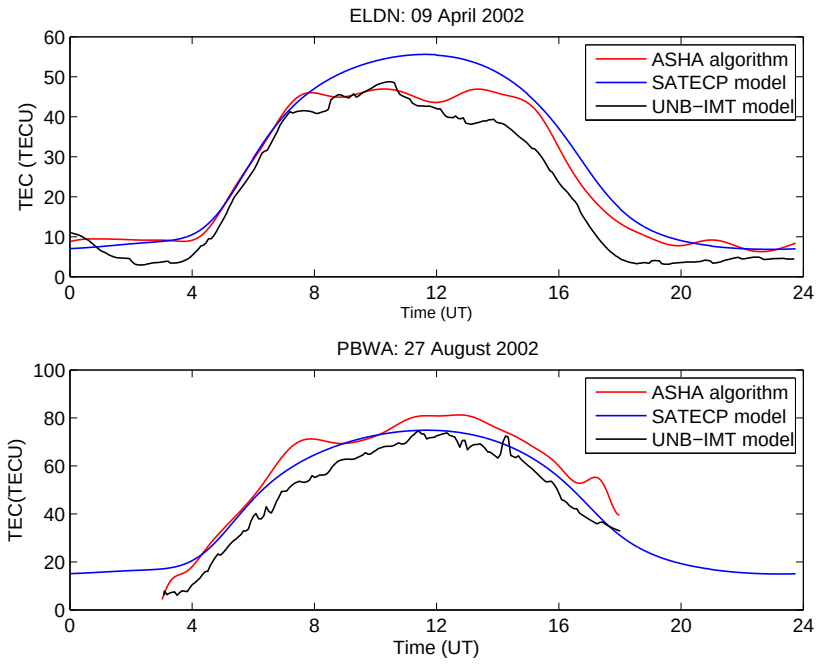


Figure 6.7: Comparison of TEC values generated by the SATECP model, the UNB-IMT model and the ASHA algorithm for ELDN and PBWA.

When using the thin shell approach, the major error sources stem from the mapping function that geometrically transforms slant TEC to vertical TEC (Komjathy, 1997). The SATECP model was developed by using ASHA-derived TEC as the target value and is thus expected to agree more with ASHA than UNB-IMT. In general, however, the three models or techniques compare well, both seasonally and diurnally. These comparative sample results reveal that the SATECP model may be used as the primary supplementary/complementary option for both the adapted UNB-IMT model and the developed ASHA algorithm over Southern Africa.

6.5.2 SATECP model validation with ionosonde data

As previously mentioned, the ionosonde measures the bottomside ionospheric variations and fits the measurement to a modelled topside when producing the electron density profile. Nevertheless, ITEC is expected to be less than the TEC obtained from any algorithm or model designed to derive or model TEC from GNSS signal information due to the altitude of operation of the two sets of instruments, which leads to a greater contribution of the plasmaspheric to GPS/GNSS-derived TEC. However, care should be taken since this fact (PTEC<VTEC) can

hide estimation errors of either GPS TEC or ITEC.

Fig. 6.8(a-c) shows the 1-minute reconstructed TEC variability by the SATECP model over the three ionosonde stations for 19-21 September 2004. Superimposed on these plots is ITEC at intervals of 30 minutes for the corresponding days. While this method is definitely not ideal for validating a physical model due to the presence of a modelled topside profile within the ITEC, it provides an idea of the general diurnal quantitative measure of TEC, since the plasmasphere is known to contribute some electron content to TEC derived from GPS (e.g. Breed et al., 1997; Belehaki et al., 2004; Mosert et al., 2007). In South Africa first comparisons between GPS TEC and ITEC over GRTN using data for four months (March - June, 2005) reported an average difference of ~ 3.57 TECU (McKinnell et al., 2007), which accounts for part of the plasmaspheric contribution. Note that both ITEC and GPS TEC contain estimation errors arising from algorithms used to derive their respective values. As expected ITEC<GPS TEC over LVLE and MDBO as shown in figs. 6.8(a-b). Over GRTN (fig. 6.8c) a similar feature is observed for 19 September 2004 and most of the time for the remaining days apart from periods when ionisation was probably high. This is not feasible physically speaking, unless there are physical processes going on within the ionosphere. This is not likely since there is no agreement with ionosonde data at other ionospheric stations. Therefore the explanation for ITEC being greater than TEC estimated by SATECP model could be the following;

- The modelled topside introduces excess electron content into the bottomside derived component, or
- The SATECP model is underpredicting TEC over GRTN for these particular times, or
- A combination of the above suggested factors.

The suggested options are all difficult to verify since there is no real GPS data for this period over GRTN as the GPS receiver co-located with the ionosonde was installed in early 2005. To further investigate this event, TEC variability from all three sources (GPS TEC, ITEC and SATECP model) for a similar period during a year that also had a decreasing trend in solar activity levels (19 - 21 September 2005) was plotted as shown in fig. 6.9. In fig. 6.9 GPS and predicted TEC are plotted at 1-minute intervals while the ITEC is superimposed at 15-minute intervals. During approximately the same period when ITEC was greater than the predicted TEC for 19-21 September 2004, a similar result is observed for 19-21 September 2005. Additionally, the SATECP model underpredicts GPS TEC during this period and the foF2 variability does not show significant changes, an indication that there were no abnormal physical processes occuring within the ionosphere. While this increases the probability that the SATECP model may have underpredicted what could have been GPS TEC over GRTN during $\sim$ 10h00 - 13h30 UT (20 September 2004) and $\sim$ 09h00 - 12h00 UT (21 September

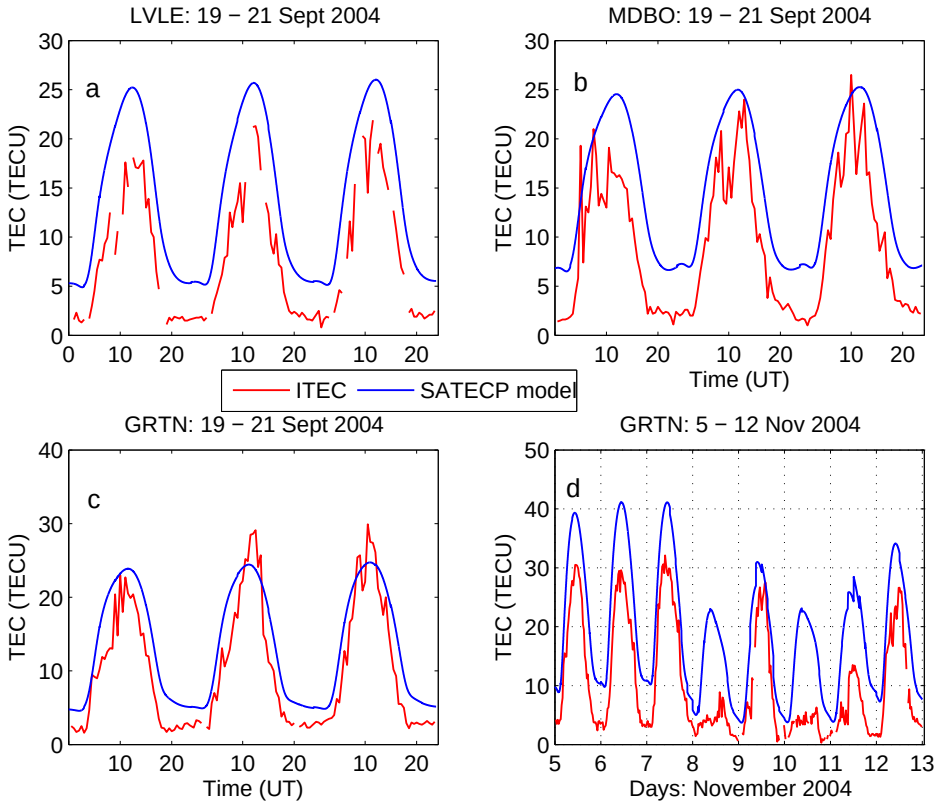


Figure 6.8: Verification of the SATECP model with ionosonde data over the three South African ionosonde stations for 19-21 September 2004 (a-c) and for GRTN during the storm of 5-12 November 2004.

2004) as is the case for roughly the same time on 20-21 September 2005, the quantification of the error introduced in ITEC as a result of the modelled topside contribution remains an open question for investigation. In fig. 6.8d the capability of the SATECP model to generalise ionospheric behaviour during storm conditions is demonstrated. A TEC decrease on 8, 10 and 11 November 2004 associated with negative Dst values using ionosonde and GPS data over GRTN and SUTH respectively, was reported by Opperman (2008). This general depletion of plasma density as observed in GPS TEC data over SUTH was reproduced by means of modelling over GRTN (where no GPS data exists for this period), with the anticipated result that ITEC is less than predicted TEC during the storm period. This example demonstrates the reliability of the SATECP model for inferring diurnal ionospheric TEC behaviour during both quiet and disturbed conditions. More details about the regional investigation of ionospheric behaviour during magnetic storms using the developed SATECP model are presented in the next chapter.

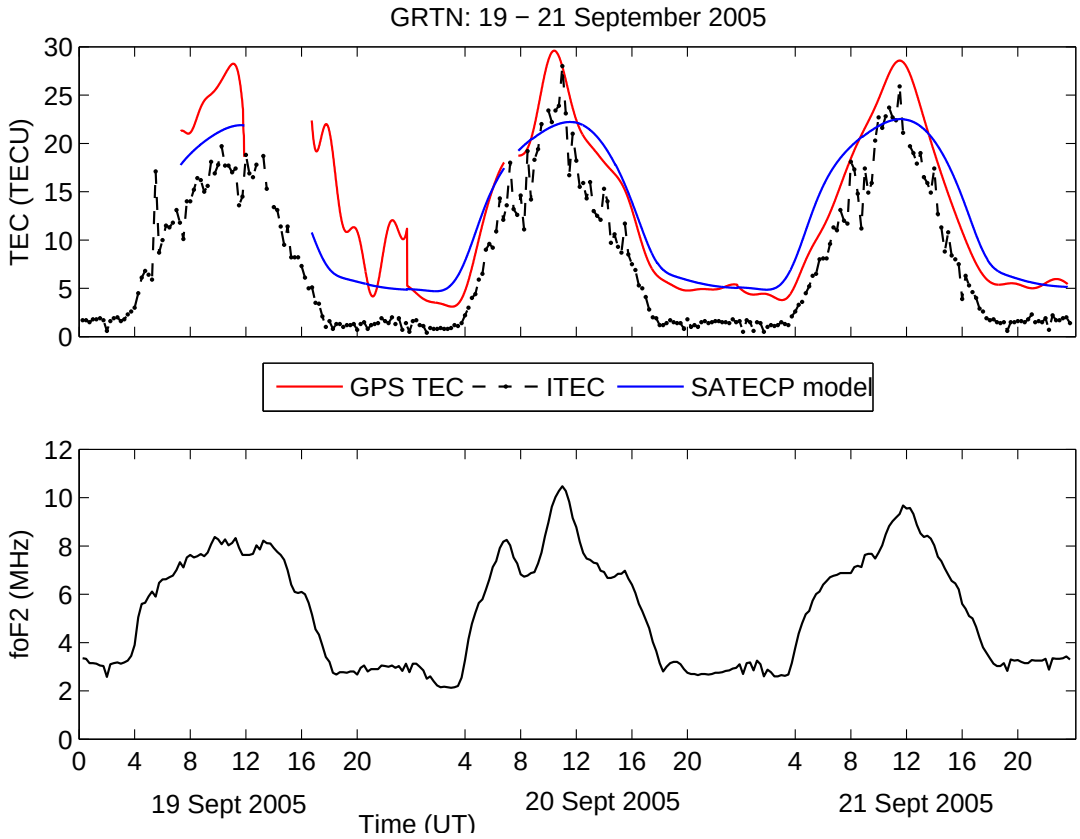


Figure 6.9: Comparison of GPS TEC, ITEC and SATECP model data over GRTN for 19-21 September 2005. The foF2 variability for the same period is also shown.

6.6 Towards reconstructed regional TEC maps

The geographical location encompassing South Africa was divided into a grid of latitude and longitude bins of dimensions $2^\circ \times 2.75^\circ$ for the generation of predictions.
After the predictions of TEC at the specified points on the grids, techniques that provide other artificially generated TEC data such as interpolation and extrapolation are required to fill in the data gaps in order to produce final smoothed TEC maps for the region of interest (Yilmaz et al., 2009). Once the model has been developed, the bin resolution can be easily varied, depending on user requirements.

6.6.1 Extrapolation and interpolation

As mentioned, predicted TEC values can be produced by the SATECP model in geographical bins of $2^\circ \times 2.75^\circ$ in the latitude-longitude space representing South Africa. However, due to the location of stations shown in fig. 6.1, extrapolation and interpolation techniques are needed to generate smoothed TEC maps for any time considered. These techniques are necessary due to the sparsity of the GPS receivers. Interpolated and extrapolated TEC values along with

predicted/computed TEC values make it relatively easy to construct ionospheric maps and analyse the temporal behaviour of TEC distribution regionally or globally (Komjathy, 1997; Yilmaz et al., 2009).

6.6.1.1 Extrapolation

Fig. 6.1 shows that the GPS stations used in the model development fall within South Africa. Outside the coverage area, the SATECP model is used to extrapolate TEC values. The model's extrapolation capability is based primarily on results obtained from single-station studies where it was estimated that the NN model developed at a particular GPS receiver station could be used to predict TEC variations at different locations within a 3-degree latitudinal separation. For example, at $\sim 1^\circ$ and 2° latitudinal separation, NN models have given maxima correlation coefficients of ~ 0.9703 and 0.905 for March equinox (19 March 2003) and June solstice (22 June 2002) respectively (Habarulema et al., 2009a). However, unlike the SATECP model, the development of single-station models do not take into account the dependance of the TEC parameter on geographical location, as described in chapter 4. Other than for single-station approximations, the extrapolation capabilities of the developed model are uncertain. Although the inclusion of the dependance of the TEC parameter on the geographic location improves single-station prediction results, NNs are known to not extrapolate well outside the input range and space.

6.6.1.2 Interpolation of TEC values

Between the bins of $2^\circ \times 2.75^\circ$ latitude/longitude, a MatLab-based biharmonic spline interpolation method in 2-dimensions is used to generate TEC values at intervals of 0.1° in both latitudinal and longitudinal space. This interpolation method uses Green functions which were originally applied in solving inhomogeneous differential equations with specified boundary conditions (http://mathworld.wolfram.com/GreensFunction.html). The description of the biharmonic spline interpolation method below relies heavily on the text of Sandwell (1987). For a 1-dimensional space, the point force Green function for the spline should satisfy the biharmonic equation

$$\frac{d^4\phi}{dx^4} = 6\delta(x) \qquad \text{with solution} \qquad \phi(x) = |x|^3 \tag{6.2}$$

The application of the Green function to interpolate N data points gives

$$\frac{d^4w}{dx^4} = \sum_{j=1}^{N} 6\alpha_j \delta(x - x_j) \tag{6.3}$$

$$w(x_i) = w_i \tag{6.4}$$

if the weight, w_i is located at point x_i in a 1-dimensional space.
In a 2-dimensional space, to interpolate N data points, the problem becomes

$$\Delta^4 w(x) = \sum_{j=1}^{N} \alpha_j \delta(x - x_j) \tag{6.5}$$

$$w(x_i) = w_i \tag{6.6}$$

The general solution to the above equations is

$$w(x) = \sum_{j=1}^{N} \alpha_j \phi_2(x - x_j) \tag{6.7}$$

where Δ^4 is the biharmonic operator, $w(x)$ is the biharmonic function, ϕ_2 is the biharmonic Green function in a 2-dimensional space, α_j is the strength of each point force and x is the position in a 2-dimensional space.
For all the reconstructed TEC maps presented in this thesis, a uniform interpolation space of 0.1° was applied both in latitude and longitude. It is possible to set the interpolation interval to less than this value, although the code takes more time for mathematical computation of the interpolated TEC values.

6.7 Summary

This chapter has described the steps taken to develop and verify the SATECP model over 8 GPS receiver stations distributed non-uniformly across the region of interest. It also presented the validation of the developed model with ionosonde data and TEC derived from the UNB-IMT model. Details were given on how TEC maps can be reconstructed by using predicted TEC values along with the brief explanation of the extrapolation and interpolation techniques for smoothing the final TEC maps. Additional results (including generated smoothed TEC maps using the techniques discussed in this chapter and potential applications of the SATECP model) are presented in the next chapter.

Chapter 7

Results and applications of the SATECP model

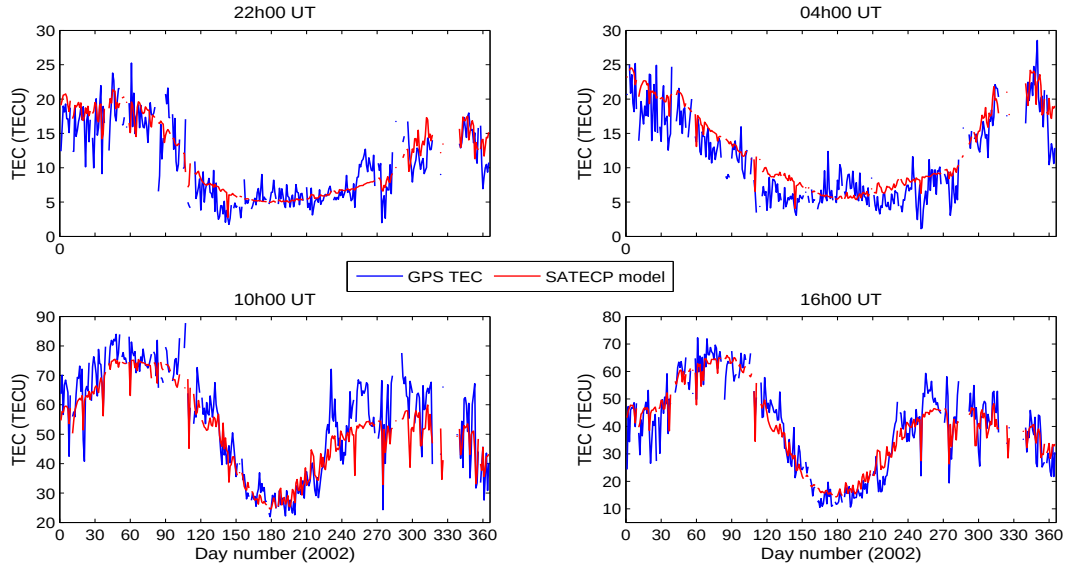
In this chapter some results and potential applications of the SATECP model are presented. Results cover data analysis for diurnal, seasonal, latitudinal and solar activity TEC predictions. To demonstrate the latitudinal TEC predictions, regional TEC maps are reconstructed using TEC data generated by the SATECP model along with the interpolation technique.

7.1 Results

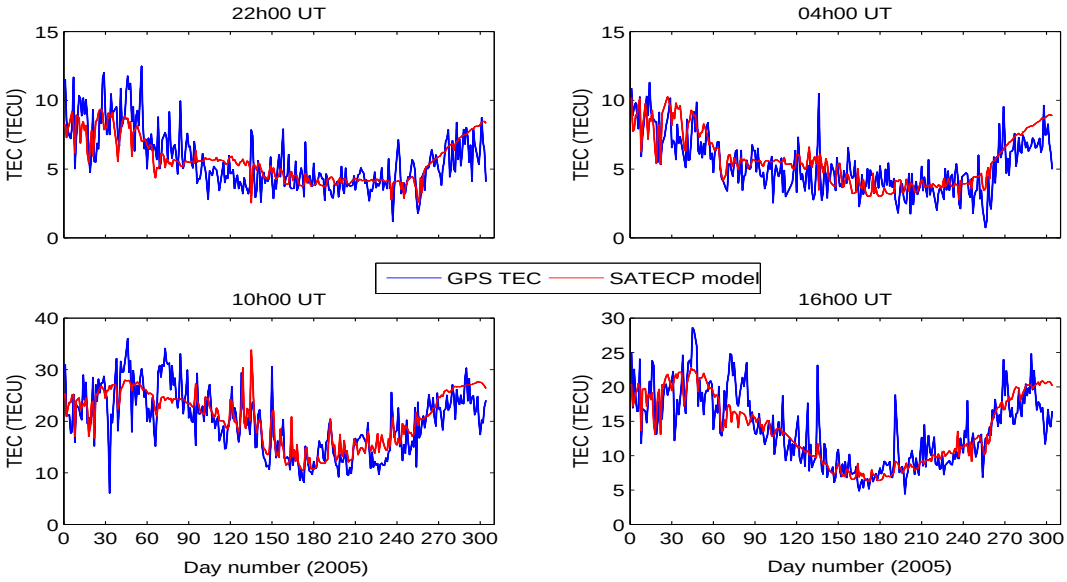
Results are presented in a form which takes into account some of the original conditions to be satisfied by the newly-developed TEC prediction model. Examples of diurnal TEC predictions were presented in the previous chapter, while the seasonal performance of the SATECP model will be demonstrated in this chapter, taking into account the solar cycle TEC prediction results.

7.1.1 During different solar activity periods

As discussed earlier, the sun is a major driver of the solar terrestrial processes through the emission of ultraviolet radiations that ionise the different layers of the atmosphere. During high solar activity, ionisation is greater within the ionosphere and hence TEC values are higher. The reverse is true during the low solar activity periods. Thus any theoretical or empirical model designed to generate TEC predictions must exhibit this property. The assessment of the SATECP model performance during different solar activity periods is shown in fig. 7.1 for the years 2002 (near solar maximum) and 2005 (solar minimum) over the verification station located at CPTN.



(a) An example of model's results near solar maximum



(b) An example of model's results near solar minimum

Figure 7.1: An example of the SATECP model's performance near solar maximum (2002) and solar minimum (2005) at local midnight, sunrise, midday and sunset over CPTN.

Prediction results for 2002 and 2005 at all considered hours, display TEC dependence on seasonal variability correctly. TEC is expected to be high during autumn and lowest during winter, an aspect that is clearly evident in SATECP model predictions. Fig. 7.2 shows the performance of the SATECP model during the low solar activity period (2007) over SUTH.

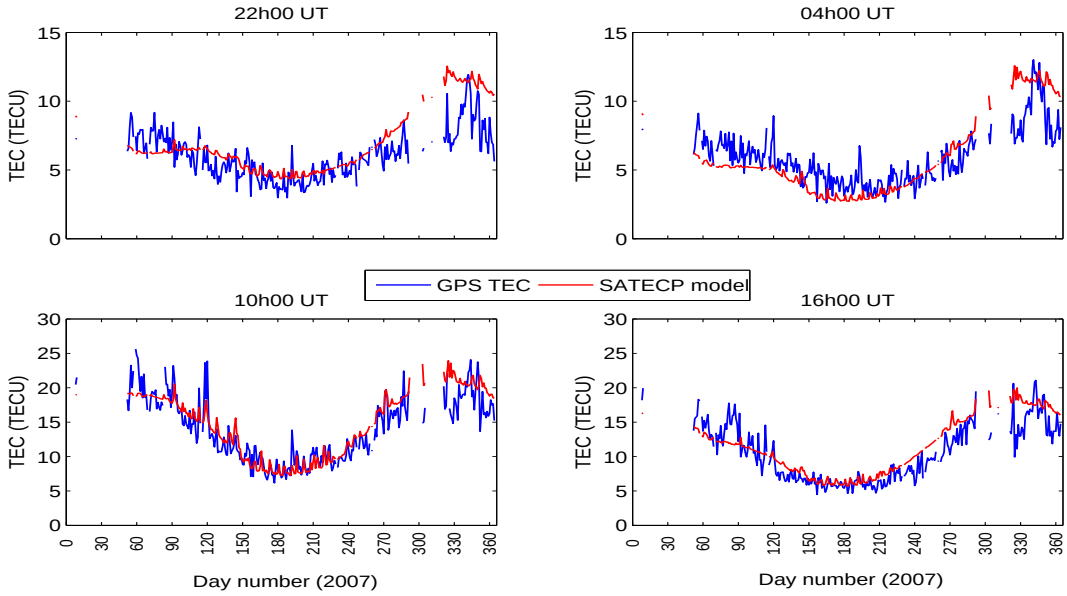


Figure 7.2: SATECP model's performance during 2007 at local midnight, sunrise, midday and sunset over SUTH.

Given that the SATECP model development utilised data for 2000-2006, this example demonstrates the capability of this data-driven model to predict unseen data in both space and time. More discussion on TEC extrapolation is found in the next section.

7.1.2 Artificial latitudinal TEC simulation

The ionosphere varies considerably with respect to the point on Earth where measurements are taken as a result of the solar zenith angle variation (McNamara, 1991). Fig. 7.3 shows an example of prediction-based TEC maps for Southern Africa at local midnight (22h00 UT), sunset (04h00 UT), midday (10h00 UT) and sunrise (16h00 UT) for day 79 in 2002. In this figure TEC is high at midday as a result of enhanced ionisation produced by the extreme ultraviolet radiation from the sun. In the Southern Africa region, TEC can be expected to increase towards the low midlatitudes, as featured in the artificial simulation results especially at 16h00 UT. At sunrise (04h00 UT) it appears that the SATECP model demonstrates the well-known phenomenon that the sun rises in the east and sets in the west. Higher TEC values are observed on the east side of South Africa at sunrise. This can be expected, since a well-trained NN model would rarely miss such important behaviour in the data. Outside the region covered by the GPS receiver distribution, the SATECP model was used to extrapolate and interpolation was later performed to generate smooth TEC maps. It is difficult to assess the general performance and extrapolation capability of the SATECP model, since there is only artificially generated TEC data available for the extrapolated sites. Examples in which reconstructed TEC and GPS TEC maps are compared, will be shown later in this chapter.

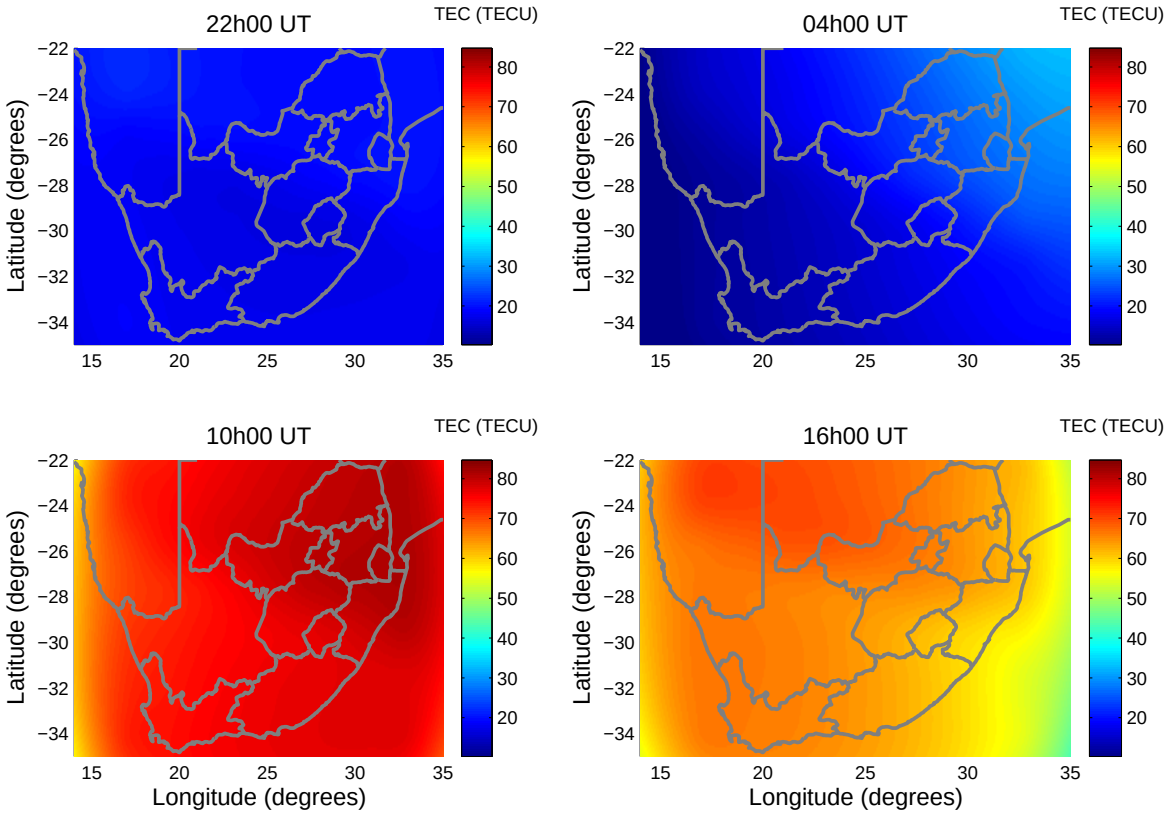


Figure 7.3: Reconstructed TEC variations over Southern Africa at local mid-night, sun-rise, mid-day and sunset for 20 March 2002. TEC values are predicted by the SATECP model.

The SATECP model in conjuction with interpolation and extrapolation methods, generate smooth TEC maps for Southern Africa. Related results using multilayer perceptrons (MLPs) for regional TEC mapping in Europe were reported by Yilmaz et al. (2009).

7.2 Potential applications of SATECP model

As mentioned in chapter 1, TEC models may be used for real-time and post-processing applications. The developed SATECP model is applicable for post-processing. It can, however, be improved for real-time applications, especially if more data becomes available with the possibility of real-time TEC forecasting.

7.2.1 Recovery of lost data through predictions

Due to mainly logistical limitations, infrastructure to record data measurements may not be available at each and every desired location, both locally and globally. Practically, however, even where equipment exists, there are some data gaps (i.e data loss) due to many factors, such

as loss of satellite signal information, software and hardware problems, equipment shutdown for maintenance purposes, etc. As a result, extrapolation and interpolation techniques are required to fill in data gaps in order to generate a complete picture of the regional and global ionospheric behaviour. Rather than performing single site extrapolation and interpolation at each receiver station affected, comprehensive models that use these techniques should be employed. By using the developed TEC prediction model, one can get an idea of TEC variation over any location within South Africa. This method can be expanded to cover the entire African continent where infrastructure presence is still scarce.

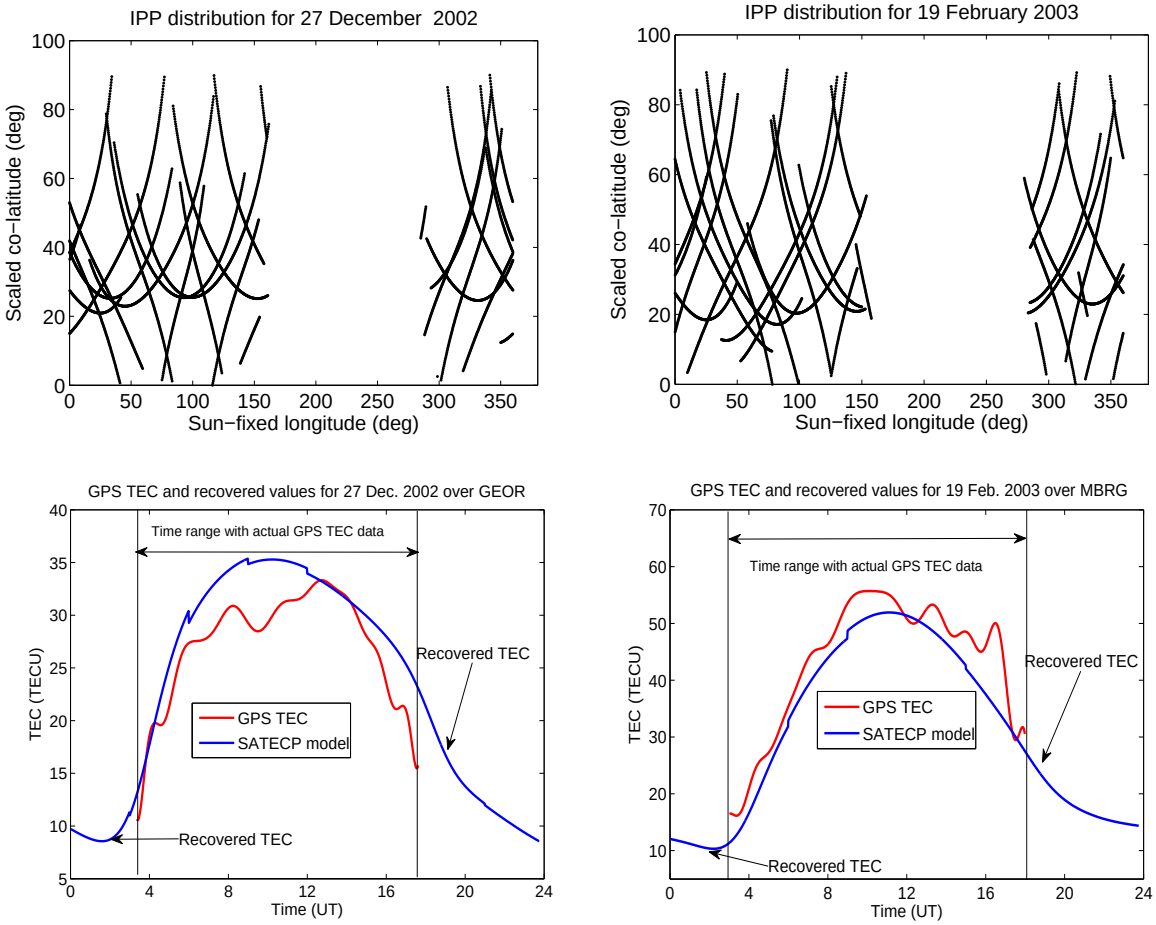


Figure 7.4: IPP distribution, GPS TEC and recovered TEC values for 27 December 2002 and 19 February 2003 over GEOR and MBRG respectively.

Fig. 7.4 shows IPP distribution, GPS TEC and recovered TEC values (through predictions) for days which originally had incomplete datasets. This application is important where particular physical aspects are being investigated in areas where instruments failed. This makes it possible to go back in time in order to artifically recover data by means of prediction.

7.2.2 Extrapolation and forecasting of TEC

As mentioned before, the extrapolation capability of the model outside its input space is not easy to determine, since predictions are artificially generated and there are no real measurements for comparative purposes. A NN with extrapolation capabilities would generate TEC predictions over specific locations and then interpolate missing values enhancing our understanding of the ionospheric TEC dynamics over a wide geographical area as well as our TEC mapping skills (Yilmaz et al., 2009; Habarulema et al., 2010). However, extrapolating and forecasting ionospheric time series of ionospheric parameters is a complex task because of sudden and dramatic changes in ionospheric behaviour, especially during magnetic storms and other ionospheric disturbances (e.g. Joseyln, 1995). Here the SATECP model is used to extrapolate TEC values during magnetic storms both in space and time. However, NNs provide the best average solution to a given dataset and the inputs during severe magnetic storms fall outside this average. Temporal forecasting using data outside the covered range was also performed. Fig. 7.5 shows areas/locations where spatial extrapolations and temporal forecasting have been performed using the developed regional SATECP model. For spatial extrapolation, two arbi-

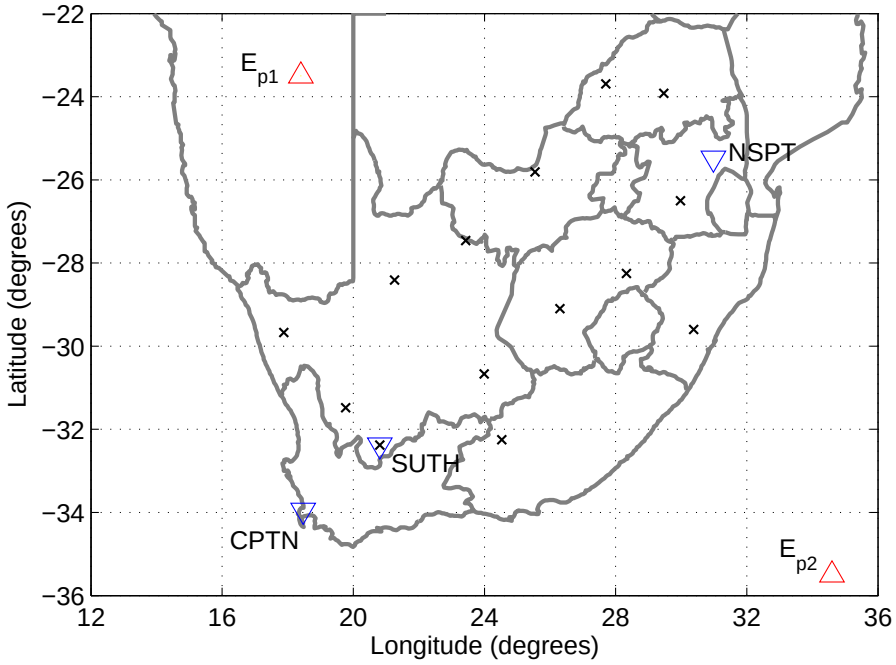


Figure 7.5: Map showing geographical locations of GPS receiver stations used in the development of SATECP model (x), spatial extrapolation sites, E_{p1} and E_{p2} (△) and verification stations (▽) for storm nature identification purposes.

trary points (in Namibia (23.5°S, 18.4°E) and over the Indian Ocean (35.5°S, 34.6°E)) which fall outside the input space were chosen and hourly TEC over these places predicted using the SATECP model. They were then compared with the IRI-2007 model during the magnetic storm

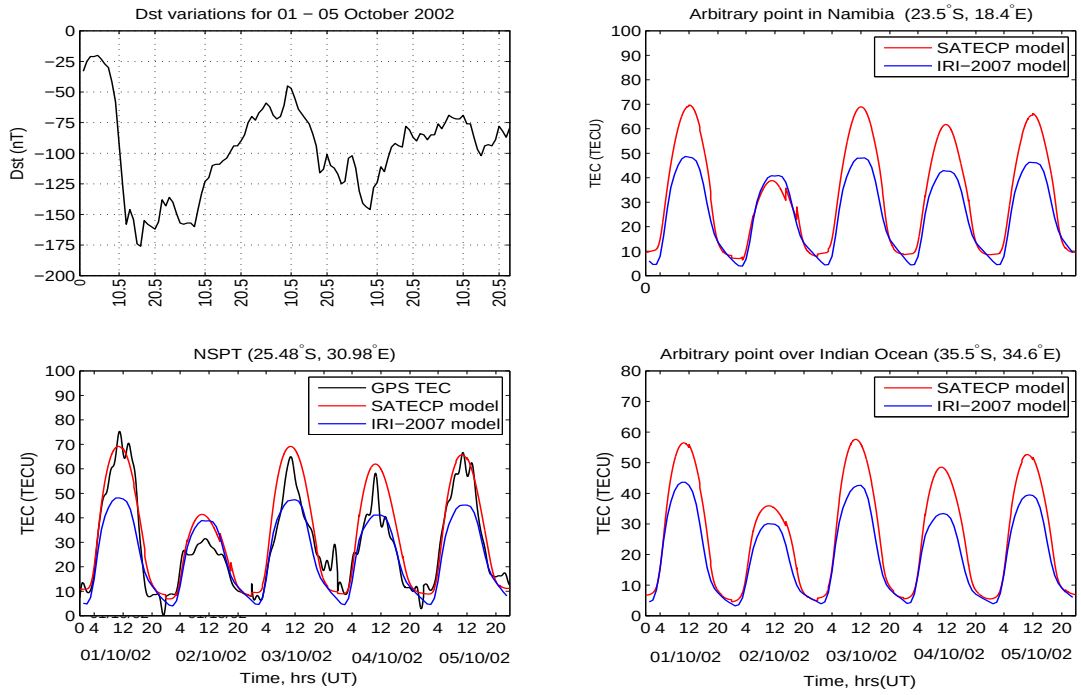
periods of 01-05 October 2002 and 23 - 26 August 2005. The main reason for comparison with the IRI-2007 model is the verification of the agreement between the global and regional models in identifying storm variational patterns, not necessarily the accuracy. The validity of the predictions in terms of storm dynamics identification, over CPTN and NSPT receiver stations are presented in the results shown in fig. 7.6. Table 7.1 shows the geographical location of stations used in both spatial extrapolation and temporal forecasting, along with the investigated time periods.

Table 7.1: Location of the stations used to verify the SATECP model and the periods for which TEC predictions were generated for spatial extrapolation and temporal forecasting.

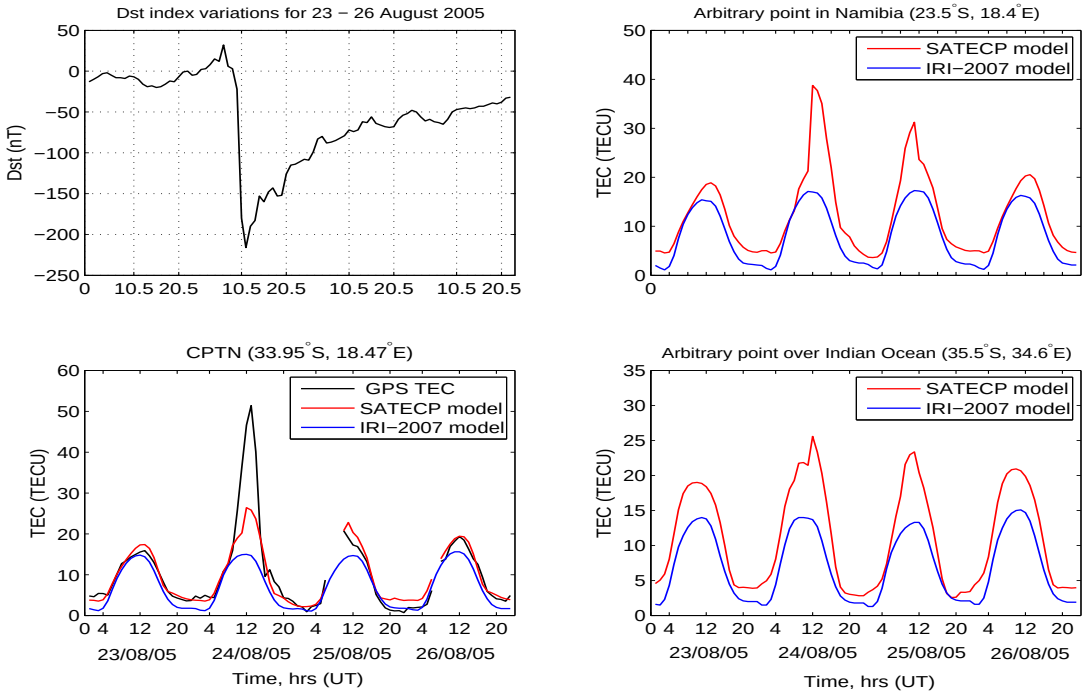
Station	Station code	Latitude (° S)	Longitude (° E)	Investigated period
Cape Town	CPTN	33.95	18.47	23-26 August 2005
Nelspruit	NSPT	25.48	30.98	01-05 October 2002
Sutherland	SUTH	32.38	20.81	22-25 May 2007
				19-22 November 2007
				Equinoxes and solstices
In Namibia	E_{p1}	23.5	18.4	23-25 August 2005
				01-05 October 2002
Over Indian Ocean	E_{p2}	35.5	34.6	23-25 August 2005
				01-05 October 2002

Fig. 7.6(a) shows a well-pronounced negative storm effect on 02 October 2002 over NSPT. The SATECP model correctly identifies the TEC variational pattern, but overpredicts TEC values between $\sim$ 08h00-16h00 UT. For the other days, the SATECP model performs relatively well in predicting TEC compared to the IRI-2007 model. Over the arbitrarily chosen points the SATECP model reproduces a negative storm effect. Higher TEC values are predicted for Namibia compared to values for the Indian Ocean. Although verification of this is a problem, due to the absence of real data, this difference is to be expected as a result of the dependence of the TEC parameter on the latitudinal position of locations used to estimate TEC predictions. Fig. 7.6(b) shows a positive storm effect over CPTN on 24 August, 2005 when TEC values were greater compared to values for days before and after the storms. The SATECP model identifies this effect correctly, but underpredicts TEC in the time range $\sim$ 10h00 - 15h00 UT. Over the arbitrarily chosen points this positive storm effect is also well identified on the same day. However, given that the model understimates TEC over CPTN during the stormy day, the overall extrapolation accuracy of the chosen points cannot be easily ascertained. In all cases the SATECP model predicts higher TEC values compared to the IRI-2007 model. This is to be expected due to the reasons discussed in the previous chapters.

For temporal forecasting fig. 7.7(a) shows TEC predictions during the magnetic storms of 22-25 May and 19-22 November 2007. The year 2007 was a period of low solar activity and therefore no major storms were observed. There appears to be no significant change in diurnal TEC

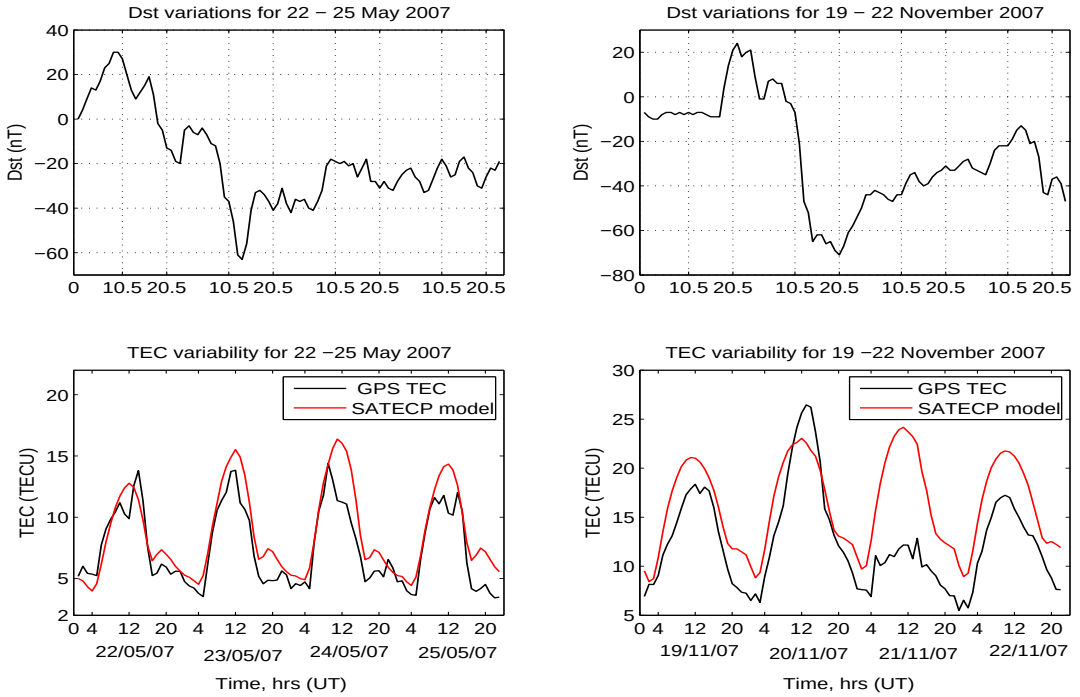


(a) NSPT and the arbitrary sites: 01-05 October 2002

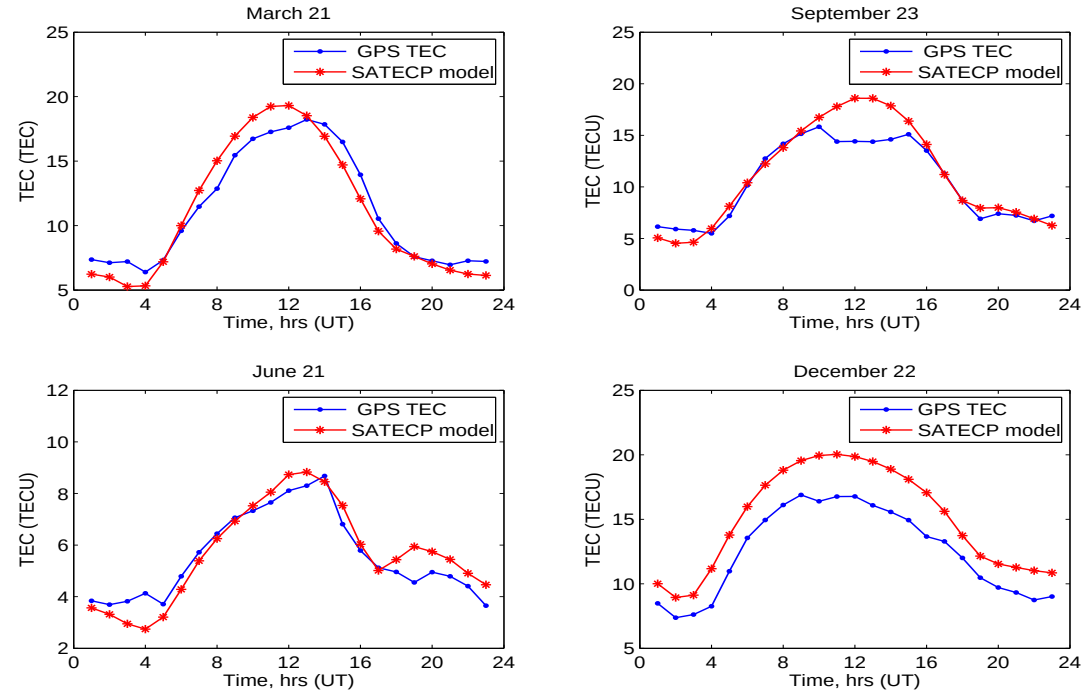


(b) CPTN and the arbitrary sites: 23-26 August 2005

Figure 7.6: Dst index variations, GPS TEC and predictions generated by SATECP and IRI-2007 models along with the respective TEC predictions over arbitrarily chosen points in Namibia and the Indian Ocean (a): NSPT: 01-05 October 2002 and (b): CPTN: 23-26 August 2005.



(a) For storm periods over SUTH and arbitrary points: 22-25 May and 19-22 November 2007.



(b) TEC variations for equinoxes and solstices in 2007 over SUTH

Figure 7.7: Forecasted hourly GPS TEC and the corresponding predictions for (a): storms of 22-25 May and 19-22 November 2007 over SUTH and arbitrary sites, and (b): days representing equinoxes and solstices over SUTH in 2007.

variability as a result of the weak storm of 22 - 25 May 2007. For this period there are no major differences between the observed and predicted TEC. During the storm period of 19 - 22 November 2007, GPS TEC observations show a slight positive storm phase on 20 November 2007 followed by a negative storm phase on 21 November. While the SATECP model tries to predict the positive storm effect, it generates a completely negative result for the negative storm effect. This is a typical example of the limitation or failure of the SATECP model to extract temporal values for magnetic storms. In fig. 7.7(b) the forecasted diurnal TEC values for March 21, September 23, June 21 and 22 December representing, the autumn and spring equinoxes, and winter and summer solstices in 2007, are shown. The accuracy of the prediction/forecasted results are highly comparable with results obtained when the verification dataset used falls within the training dataset range. This demonstrates that for quiet periods, TEC forecasting can be done with some degree of accuracy. During the December solstice the prediction accuracy reduces compared to other periods.

7.2.3 Investigation of magnetic storms on TEC behaviour

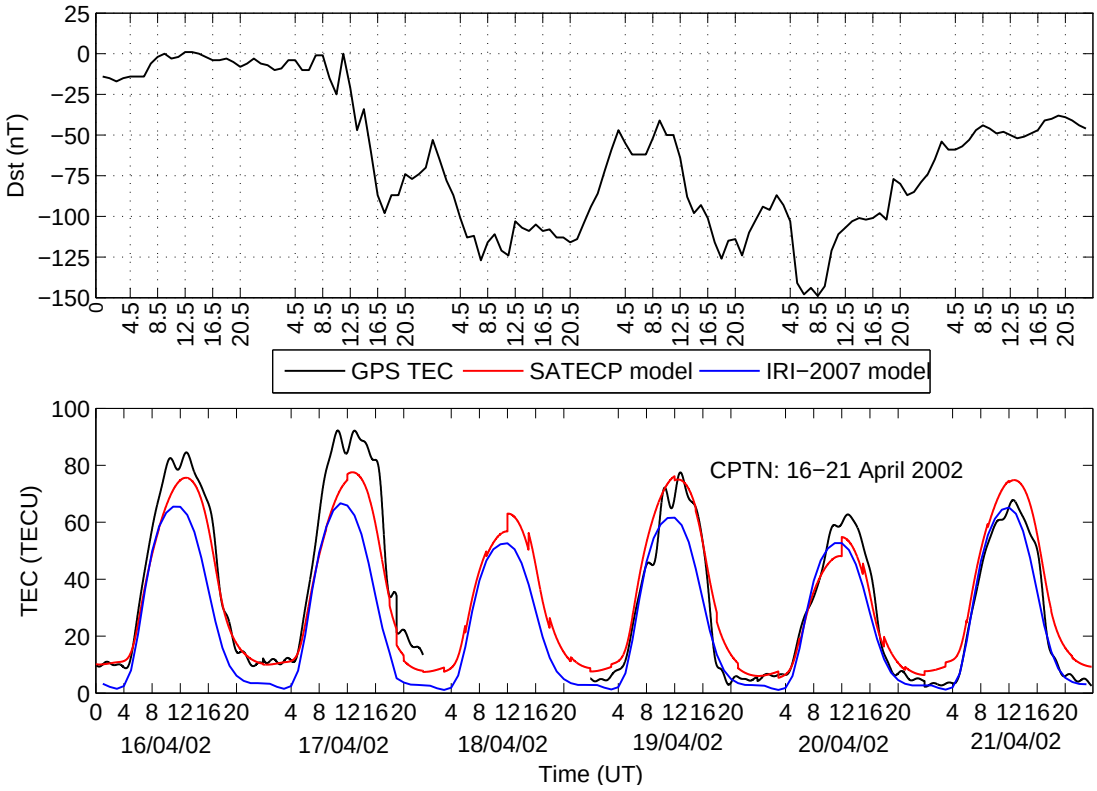
During quiet times the ionosphere tends to exhibit regular and repetitive variations diurnally and seasonally. Thus for radio wave propagation purposes, empirical models may be required to estimate corrections needed as a result of ionospheric effects (Ondoh and Marubashi, 2000). The primary source of ionospheric formation is the solar radiation. The sunspot cycle has $\sim$ 11 years. If this was the only factor responsible for ionospheric dynamics, then with the available information and developed models, the prediction of ionospheric parameters such as foF2 and TEC wouldn't be very difficult. However, at times the coupling of the solar wind - magnetosphere - ionosphere system introduces other sudden variations in TEC that alter its regular variability. The ionospheric effects due to magnetic storms are mainly characterised in terms of negative and positive effects, referring to the reduction and increase in electron density during the disturbed times compared to relatively quiet days before and after the storm duration. Over Southern African GPS receiver stations, the TEC behaviour during magnetic storms has been observed to be characterised mainly by the depletion of the electron density during the main phases of magnetic storms (Habarulema et al., 2009c, 2010). To investigate the ability of the SATECP model to characterise local TEC dynamics during geomagnetically disturbed conditions, the following storm periods were chosen;

- 16-21 April and 1-6 October 2002
- 28 October - 2 November 2003.

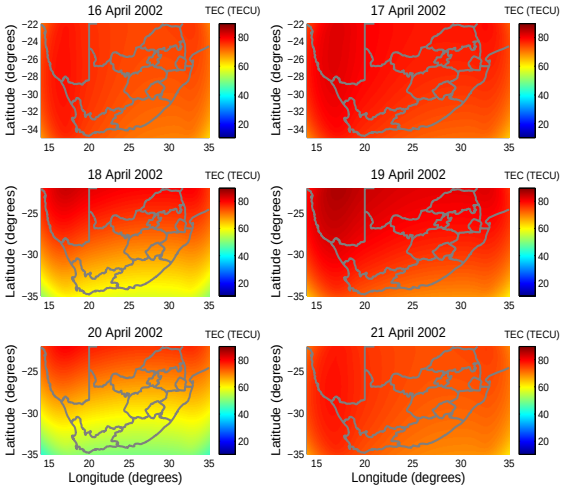
These storms were chosen on the basis of their intensity and the need to demonstrate the ability of the SATECP model to simulate regional TEC variations in the absence of sufficient historic GPS data.

7.2.3.1 Storms of 16-21 April 2002 and 1-6 October 2002

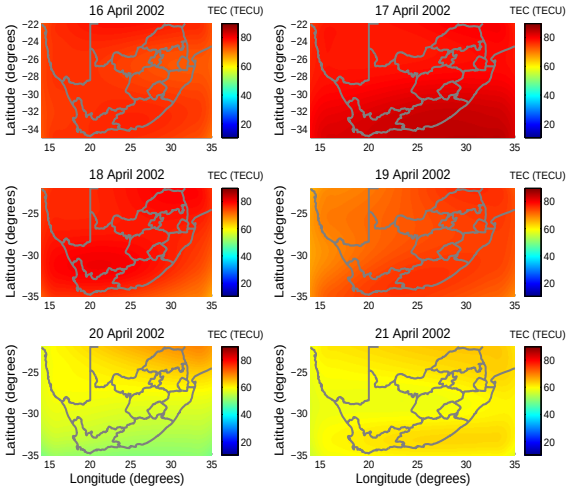
Fig. 7.8(a) shows the Dst and GPS TEC variations with the corresponding predictions by the SATECP model and the IRI-2007 model for the period of 16-21 April in 2002 over CPTN.



(a) Comparison of GPS TEC with the corresponding predictions from the SATECP and IRI-2007 models during 16-21 April 2002 over CPTN station (33.95°S, 18.47°E)



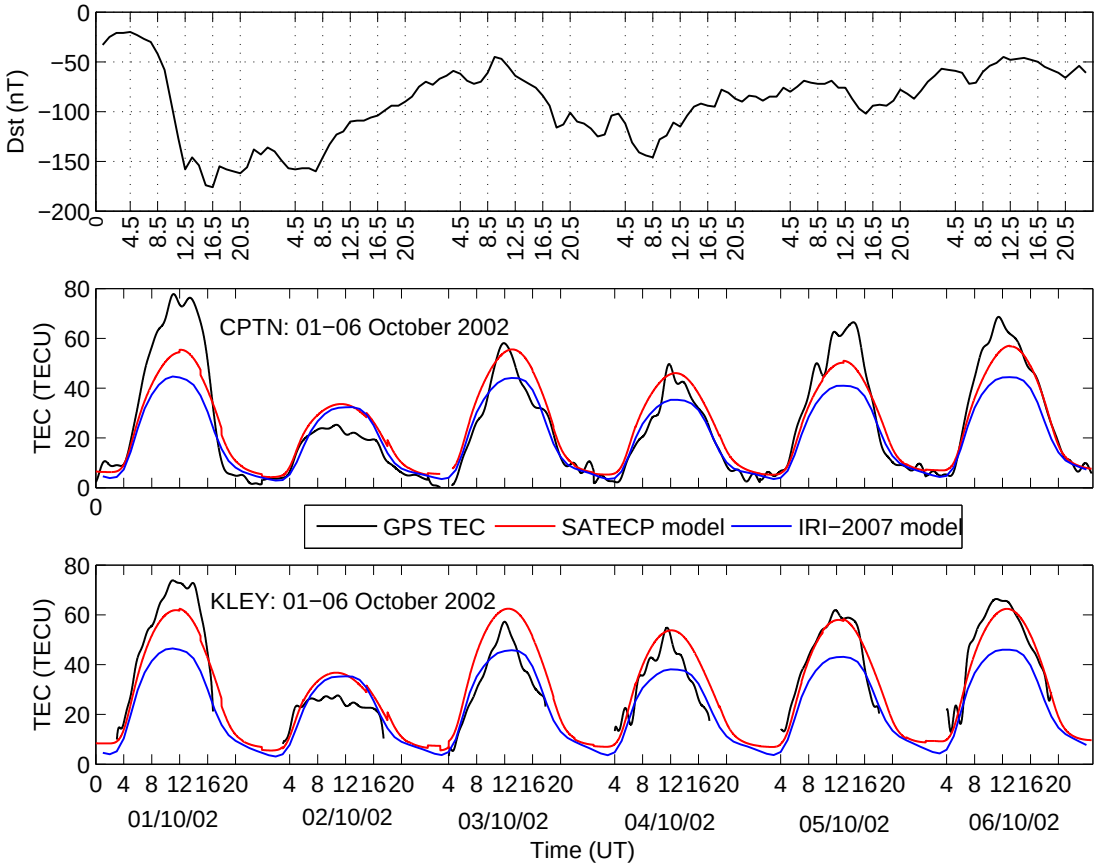
(b) SATECP model maps at 12h00 UT



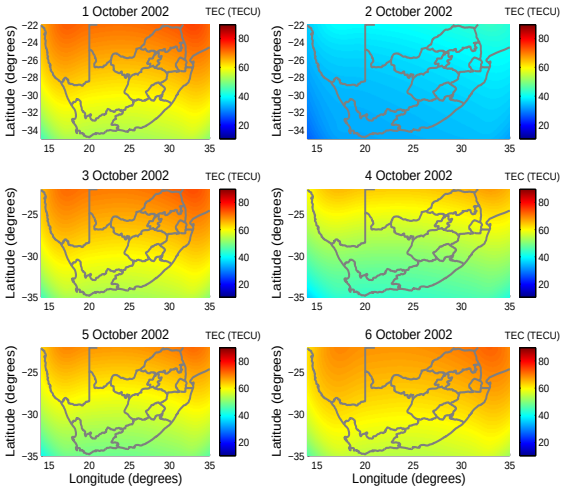
(c) GPS TEC maps at 12h00 UT

Figure 7.8: Dst and TEC dynamics during the storm of 16-21 April 2002.

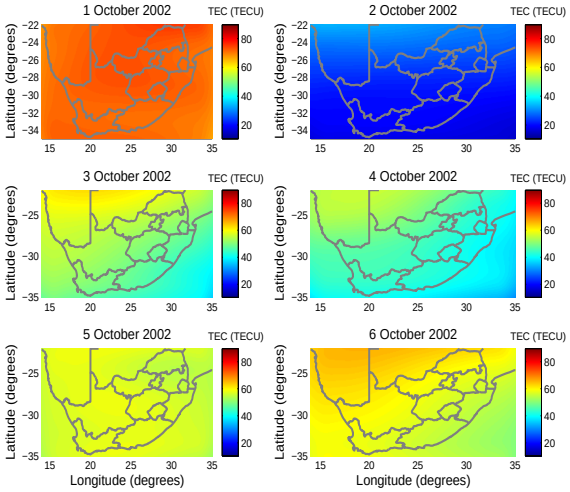
Figs. 7.8(b)-7.8(c) show the predicted TEC and actual GPS TEC maps for the entire South Africa for 16-21 April at 12h00 UT when TEC is expected to be high. Fig. 7.9 shows the same comparison for the period 1-6 October 2002.



(a) Comparison of GPS TEC with the corresponding predictions from the SATECP and IRI-2007 models during 1-6 October 2002 over CPTN and KLEY stations



(b) SATECP model maps at 12h00 UT



(c) GPS TEC maps at 12h00 UT

Figure 7.9: Dst and TEC dynamics during the storm of 1-6 October 2002.

Fig. 7.9 shows that TEC decreased significantly on a storm day (2 October) over the CPTN and KLEY stations compared to the relatively quiet days before and after the storms. Both the SATECP and IRI-2007 models identify the negative storm phase trend correctly on 2 October. The first plates of figs. 7.8(a) and 7.9(a) showing Dst variations indicate that the magnetic storm of 1-6 October was stronger than the one of 16-21 April. Both the SATECP and the IRI-2007 models are not able to predict TEC accurately during the main phase of strong magnetic storms. Fig. 7.9 demonstrates the ability of both the SATECP and IRI-2007 models to in some cases accurately capture the diurnal TEC variability for the negative storm effect during the early morning hours of 2 October especially over CPTN. For this day both models overpredict TEC during the hours $\sim$ 07h00-16h00 UT. The IRI over- or underpredicts TEC for some of the periods considered, but is close to GPS TEC in other cases, especially for quiet conditions during morning and evening hours. This was previously observed over South Africa (Habarulema et al., 2007, 2009a). The overprediction is attributed to a lack of data in the IRI for the Southern Africa region (McKinnell, 2002). Apart from 02 October, there is a general trend for the SATECP model to predict greater TEC values than those generated by the IRI-2007 model for 1-6 October 2002, although there is an under-/overprediction of GPS TEC by the SATECP model in some cases. The under-/overprediction by the SATECP model may be related to the level of data availability and the fraction of solar activity covered by the model. As pointed out in Habarulema et al. (2007, 2009a), the solar activity may not be effectively represented into the model when considering a data period that is shorter than the sunspot cycle of $\sim$ 11 years. Both the SATECP model and the IRI-2007 model underpredict TEC for what appears to be a positive storm effect observed on 17 April ($\sim$ 07h00 upwards) and 1 October ($\sim$ 06h00-17h00 UT) over CPTN. With the STORM option of the IRI activated, related results at low latitudes have been reported (e.g. Liu et al., 2004; Kutiev et al., 2007). For situations where there is an over-/underprediction for diurnal TEC variations, there is a corresponding phenomenon in predicted TEC maps compared to actual GPS TEC maps, as shown in figs. 7.8(b)-(c) and 7.9(b)-(c). These results clearly indicate that the SATECP model is capable of following TEC dynamics during magnetic storms over the entire region of interest.

Table 7.2 shows the computed RMSE and bias values between the TEC derived from GPS observations (measured TEC) and the SATECP model, along with the values generated by the IRI-2007 model during the storms of 16-21 April, 2002 and 01-06 October, 2002 over CPTN. The computed averages in this table show that the IRI-2007 model provides more accurate TEC predictions for 21 April and 02-04 October than the SATECP model. The overall average indicates that the SATECP model performs more accurately than the IRI-2007 model during the geomagnetic storms of both 16-21 April, 2002 and 1-6 October, 2002.

Table 7.2: Average RMSE values and biases (in TECU) between GPS TEC and predicted values (by the SATECP and IRI-2007 models) over CPTN (33.95°S, 18.47°E) during the storms of 16-21 April and 01-06 October, 2002.

Date	SATECP		IRI-2007	
	RMSE	Bias	RMSE	Bias
16	5.891	3.511	14.903	13.064
17	12.033	9.381	22.308	18.804
18	-	-	-	-
19	7.179	-5.251	9.006	5.131
20	5.532	2.021	7.153	4.770
21	8.479	-8.058	4.277	1.601

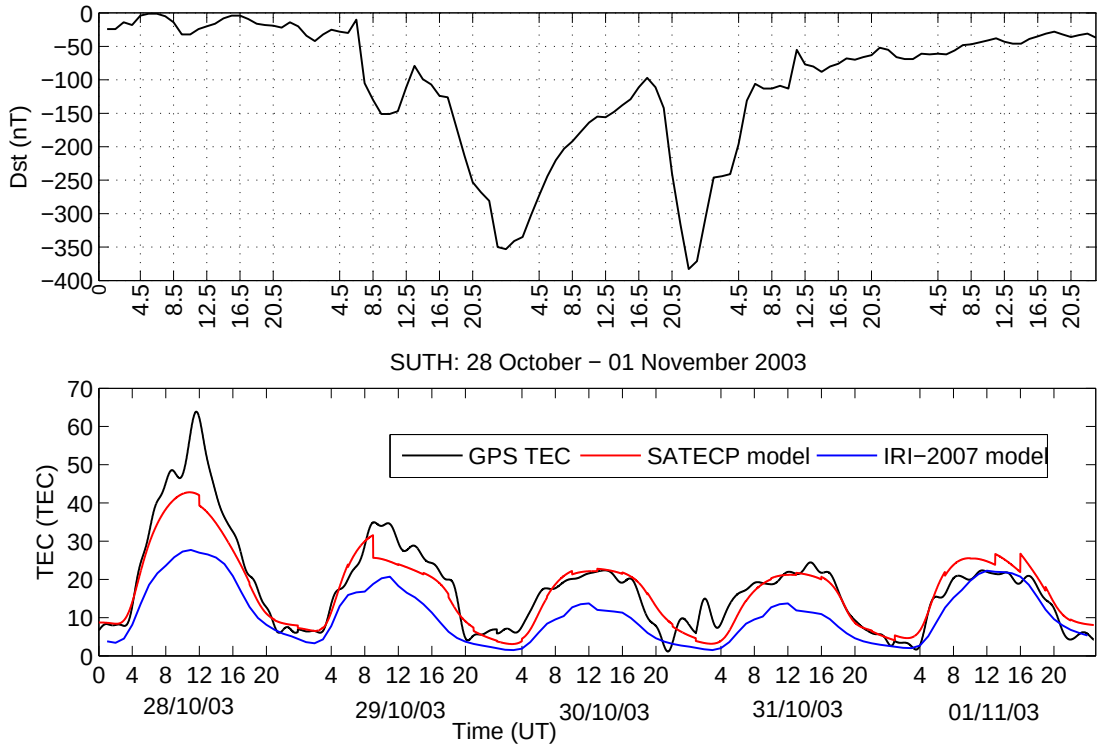
(a) CPTN: 16-21 April 2002

Date	SATECP		IRI-2007	
	RMSE	Bias	RMSE	Bias
01	12.643	6.312	17.860	11.875
02	6.232	-5.145	5.460	-3.323
03	6.750	-4.747	5.523	1.110
04	5.729	-4.961	4.165	0.574
05	7.474	3.057	11.637	7.598
06	5.530	2.459	10.997	8.003

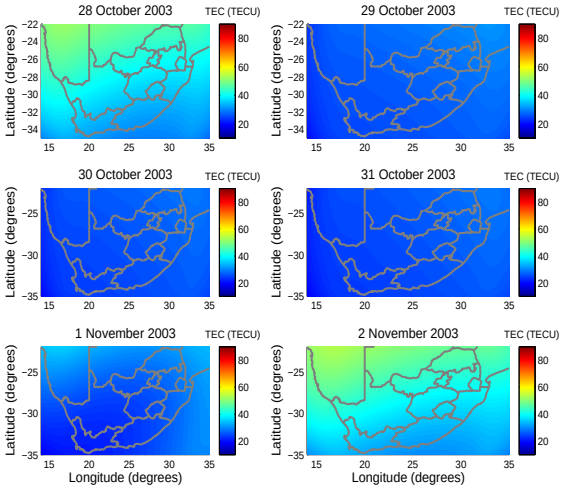
(b) CPTN: 01-06 October 2002

7.2.3.2 Storm period of 28 October to 2 November 2003

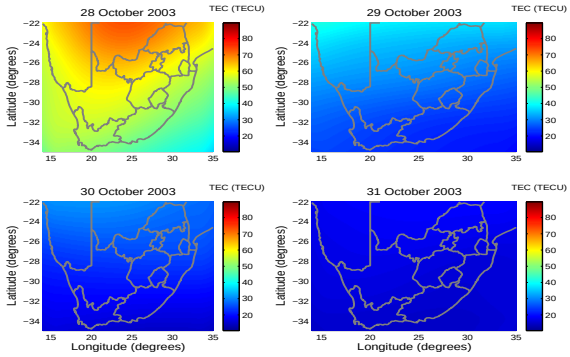
There was no validation station with complete GPS data for the geomagnetically disturbed period of 28 October-2 November, 2003. Thus the validation was also done using data from the station located at SUTH that was used in the development of the SATECP model. Fig. 7.10(a) shows the variations of Dst index and GPS-derived TEC, along with the corresponding predictions from the IRI-2007 and SATECP models. Table 7.3 presents the computed averages of RMSE and bias values between GPS TEC and the predicted values by the IRI-2007 and SATECP models. This table indicates more accurate TEC predictions by the SATECP model compared to the IRI-2007 model. It is important to mention that data for SUTH was included in the SATECP model development and therefore these results may not necessarily guarantee that the SATECP model will always provide more accurate predictions than the IRI-2007 model during magnetic storms. However, previous studies have shown that on average, locally developed NN models perform more accurately than the IRI model during quiet conditions over locations within South Africa (Habarulema et al., 2007, 2009a). Fig. 7.10(b) shows the predicted NN TEC maps at 12h00 UT over Southern Africa during the storm period of 28 Oct - 2 Nov 2003. There was not enough historic data to generate GPS TEC maps for 1 and 2 November 2003 and hence GPS TEC maps shown in fig. 7.10(c) are for 28-31 October only. The SATECP model provides the advantage of being able to infer TEC values for days without data, locations with no instruments or when the infrastructure experienced problems related to power, maintenance and other technical failures. As observed from the GPS TEC maps, SATECP model maps at 12h00 UT also show that TEC was significantly depleted over the region considered. Fig. 7.10(a) indicates an overprediction of GPS TEC by the SATECP model on 1 November over SUTH. There are also corresponding high values on the predicted TEC map for this day which shows the recovery phase of TEC and seems to be in agreement with the Dst values.



(a) An example of the storm impact on GPS TEC variability with the corresponding values of the IRI-2007 and SATECP models over SUTH (32.38°S, 20.81°E).



(b) SATECP model maps at 12h00 UT



(c) GPS TEC maps at 12h00 UT. Due to unavailability of data GPS TEC maps for 1 and 2 November are not shown.

Figure 7.10: Dst and TEC dynamics during for the Halloween storms of 28 October to 1 November 2003.

Fig. 7.11(a) shows a comparison of the SATECP model values with derived GPS TEC, IRI-2007 TEC and ionosonde TEC (ITEC) over UPTN. Both the ionosonde and GPS TEC data were not enough to draw conclusions about the ionospheric TEC dynamics during this period. However, the available data show that the ionosonde TEC is less than the GPS TEC, as expected. This

Date	SATECP model		IRI-2007 model	
	RMSE	Bias	RMSE	Bias
28/10/2003	7.197	4.182	15.775	12.589
29/10/2003	4.366	2.061	9.726	8.432
30/10/2003	3.440	-0.312	7.485	6.863
31/10/2003	3.619	2.030	8.117	7.325
01/11/2003	3.523	-3.231	2.774	1.341

Table 7.3: Average RMSE values and biases (both in TECU) between GPS TEC and predicted values (by the SATECP and IRI-2007 models) over SUTH (32.38°S, 20.81°E) during the storm period of 28 October- 01 November, 2003.

is primarily due to the altitude at which both sets of data are estimated. As mentioned earlier, the ionosonde data also serves as a validation tool for the algorithm used to derive GPS TEC.

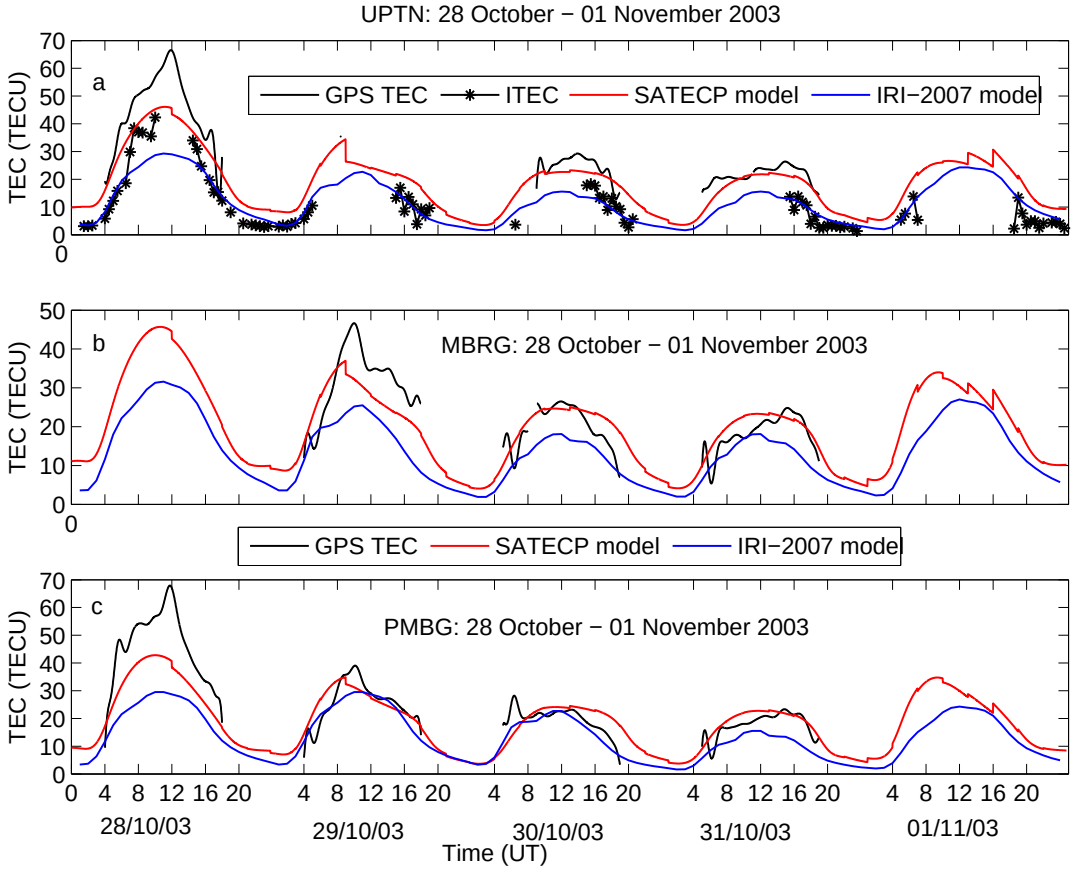


Figure 7.11: Comparison of GPS TEC with the corresponding predictions from the SATECP and IRI-2007 models during 28 October - 01 November 2003 over (a) UPTN, (b) MBRG and (c) PMBG. In the first graph, the Louisvale ITEC is also shown.

Figs. 7.11(b)-(c) show the same comparison as fig. 7.11(a), but with no ionosonde data. Fig. 7.11 shows that there are periods when there is absolutely no GPS data and so the variation

of GPS TEC over these stations is unknown. The SATECP model was used to predict TEC values for these periods as shown in figs. 7.11(a)-(c) and compared to the values generated by the IRI-2007 model. In most cases, the TEC values provided by the SATECP model are greater than the TEC values generated by the IRI-2007 model. Again, this result is to be expected since the SATECP model is GPS-based and predicts TEC to an altitude of $\sim$ 20,000 km, in contrast to the IRI-2007 that provides TEC values to a maximum altitude of $\sim$ 2000 km. Both the SATECP and IRI-2007 models underpredict GPS TEC on 28 October, 2003 and predict correctly the negative storm phase during 29-31 October, 2003, but with varying accuracy in all cases. The SATECP model does not predict TEC variation accurately during severe magnetic storms. This is also evident in fig. 7.10(a) where the data used for validation was also included for the model development. However this is not a surprising result, since the NN model provides the best average solution for a given dataset and the inputs for severe magnetic storms fall outside this average. Previous studies have shown that the NN model works relatively well with quiet times and therefore it would in future be advisable to investigate a storm time input to include within the input space, so as to provide the NN with the information over and above the average that is needed to generalise TEC patterns during magnetic storms when interrogated.

7.3 Summary

This chapter presented results and some potential applications of the developed SATECP model. These include the recovery of lost data by means of predictions and the investigation of magnetic storm behaviour over a wide geographic area for post-processing applications. It also dealt with the investigation of the capabilities and limitations of the SATECP model for extrapolation purposes in TEC modelling. Both data outside the training dataset (temporal extrapolation/forecasting) and outside the data coverage area (spatial extrapolation) were investigated. For both spatial and temporal extrapolation, predictions during quiet and disturbed conditions were generated by the SATECP model and compared to the IRI-2007 model predictions for the spatial case. Results indicate that temporal extrapolation is better during quiet periods than magnetic storm periods, while spatial extrapolation predictions identify geomagnetic storm dynamic patterns correctly. However, the accuracy is not known, since there are no real measurements for comparison analysis. Different research groups have used different types of ANNs for the forecasting of ionospheric parameters (e.g. Cander et al., 1998; Cander, 1998; Chan and Canon, 2002; Tulunay et al., 2006; Senalp et al., 2008; Yilmaz et al., 2009). To objectively understand and reach a conclusion about the limitations of NNs for the modelling/forecasting of ionospheric parameters, a comprehensive study analysing and comparing performances of different NN algorithms on a similar dataset is required. This study,

which utilised feed forward networks with a standard back propagation algorithm, reveals that forecasting TEC during magnetic storms is a difficult task, although forecasted TEC time series approximations during quiet conditions may be possible. This observation is based on the forecasted TEC for a low solar activity period when the occurrence of magnetic storms was minimal. The validity of the developed model for forecasting temporal TEC during high solar activity is also not known, but is likely to be poor since there is not enough high solar activity data incorporated in the SATECP model. The SATECP model (along with interpolation and extrapolation techniques) is capable of producing smooth TEC maps over Southern Africa. Smoothed TEC maps have also been constructed with the aid of MLP networks using GPS data for some parts of Europe (e.g. Tulunay et al., 2006; Yilmaz et al., 2009).

Chapter 8

Concluding remarks

This study explored and utilised linear and non-linear regression techniques in efforts to construct a comprehensive TEC simulation model for Southern Africa. The non-linear approach (NNs) gave more favourable results than linear regression models and was thus adopted for TEC modelling. Due to a lack of data and the difficulty to obtain data for some regions in Southern Africa, the discussion and model construction concentrated mainly on South Africa. The development of reliable Physics-based mathematical prediction and forecasting techniques for both global and regional ionospheric purposes and applications that take into account different conditions would be more suitable. This however turns out to be a very complicated task due to the fact that there are many unpredictable variables that are responsible for the ionospheric behaviour (e.g. Joseyln, 1995; Tulunay et al., 2006; Habarulema et al., 2009c). As a result data-driven (empirical and semi-empirical) approaches have been found to be more desirable, with NNs playing a crucial role, since they involve parallel processing of multiple and non-linear complex systems (e.g. Cander et al., 1998; Chan and Canon, 2002; Poole and Poole, 2002; Oyeyemi et al., 2006; Senalp et al., 2008; Tulunay et al., 2006; Yilmaz et al., 2009). A multi-station model was developed as a function of multiple inputs which were fed into the NN software for "learning" TEC variability through a mapping process. This last chapter provides a brief summary of what has been discussed in this thesis and attempts to put forward suggestions for improving on this study in future.

8.1 Summary

In South Africa the GPS infrastructure with continuous and accessible GPS data starts in year 2000. The study reported here commenced in August 2007 and hence the SATECP model was developed using data for the period 2000-2006 and verified for different solar activity levels within the limit of its validity. The forecast was done only for a low solar activity period (2007) and therefore its forecasting capability is not known for high solar activity periods. It may,

however, not perform well during the solar maximum, since high solar activity data within the model is limited. In general terms, this study was done in three distinct phases;

A feasibility study: As described in chapter 4, the main intention of this study was to establish the applicability of NN and linear regression models for modelling/predicting TEC values for locations where data was not included in the models' development. By using separate linear regression and NN models developed using data from GPS stations located at SUTH (32.38°S, 20.81°E) and SBOK (29.67°S, 17.88°E), the TEC variability was predicted for CPTN (33.95°S, 18.47°E) and UPTN (28.41°S, 21.26°E) respectively. A comparison of GPS TEC with the corresponding values obtained from IRI-2001 was also carried out. The major findings of this study were that NNs are capable of characterising the spatial and temporal variability of TEC and that the NN model in general gave more accurate TEC predictions than the IRI-2001 model. It was, however, found out that the IRI-2001 model provides accurate predictions for spring equinoxes (Habarulema et al., 2009a).

Investigating the solar wind parameter in TEC modelling: With the prospect of capturing plasmaspheric TEC variations, the solar wind parameter was included separately in the NN along with $D_n, H_r, R4$ and $A8$, as detailed in chapter 5. The considered solar wind parameters were IMF B_z, V_{sw}, N_p and P_{dyn} from both the ACE and OMNI 2 datasets. The first four parameters were taken as benchmark values for quantitatively determining solar wind effects on GPS TEC. With the inclusion of the solar wind parameter, results exhibited both improvement and degradation in prediction accuracy, both being close to the benchmark values. This gives weight to the belief that the first four parameters may be the cause of major influences on TEC variability. It is for this reason that chapter 6, which presents details of the SATECP model, does not include the solar wind parameter as one of the inputs. This study confirmed results of the feasibility study (chapter 4) with regard to spatial and temporal characterisation of TEC for a location not involved in the model's development.

Development of a regional TEC simulation model - SATECP model: Combining the feasibility studies carried out in this study, a multi-station model was developed as a function of diurnal and seasonal variations, solar and magnetic activities and the geographical location of the GPS receiver stations. First results demonstrating the possibility of using data from 10 GPS receiver stations distributed non-uniformly over South Africa for a period of utmost five years (2000-2004) were presented in Habarulema et al. (2009b). Chapter 6 is an extension of this work in that the data period in the input space was increased by 2 years from 2000-2004 to 2000-2006. Thus the temporal predictability coverage of the SATECP model is broader than the first attempt. Additionally, the number of GPS receiver stations was increased from 10 to 14 for more spatial coverage.

The final model was validated for both quiet and disturbed conditions, and also using independent data sources. The SATECP model provides expected results over MDBO and LVLE ionosonde stations and tends to underpredict TEC in some instances for the chosen verification periods for GRTN. A comparison of the SATECP model TEC values with the UNB-IMT-derived TEC indicate a similarity, both seasonally and diurnally.

While it is possible to map and improve understanding of ionospheric dynamics over a particular regional area using locally derived GPS data and models, there are some aspects of the SATECP model that require improvement which include:

- Accurately modelling TEC variability during disturbed conditions: The inability of the SATECP model to accurately capture storm TEC variability may be attributed to the presence of more quiet data compared to data for disturbed conditions. While the trend of the negative storm effect is well identified and fairly modelled, the positive storm effect is difficult to model, especially in terms of TEC magnitudes.

- Long-term temporal and spatial TEC modelling: Data availability is an important pre-requisite for effective NN modelling (e.g. Haykin, 1994; McKinnell, 2002). Therefore, attempts to model TEC variation using NN-developed models do not yield accurate results for areas and periods not effectively covered within the input space. Thus the SATECP model fails to extrapolate outside the input space (temporal modelling), especially over a long period of time. This is an expected phenomenon. Additionally, due to the data coverage within the model, comprehensive validation for both high and low solar activity periods is not possible. The SATECP model contains more data for the declining phase of the solar cycle, and according to NN applications and principles, the best results can be obtained within this limit of validity.

- Spatial data representation: TEC variations over ocean areas and outside South Africa are unknown. Additionally NNs are known not to extrapolate well outside the input space and therefore the extrapolation may be unreliable. The extrapolation capability of the SATECP model is based only on single station modelling results. Though TEC values should increase towards the lower midlatitudes, a feature that is evident in diurnal TEC predictions and smoothed TEC maps, this may not be a "good way" of generalisation. Future work should include using other data sources so that real data over the ocean and surrounding areas is included within the SATECP model.

8.2 Recommendations

This study provides insight into what can be done to predict TEC variability over a relatively large geographical area (part of Southern Africa). For quiet times the NN technique seems to

be adequate, and it also correctly identifies TEC variational patterns during disturbed conditions. Negative storm effects are particularly well identified with greater accuracy than for positive storm effects. Nevertheless, there are suggestions for improving our understanding of ionospheric prediction over the Southern Africa region and probably beyond, to cover the entire African continent.

1. Continuous ionospheric data is necessary to develop comprehensive and fully representative ionospheric models. Since instruments, such as local satellites and ionosondes, are more expensive than GPS/GNSS receivers, densification of these receivers is required to acquire TEC data in an affordable way. Apart from South Africa, other Southern Africa Development Community (SADC) countries do not have continuous GPS data from receiver stations. In cases where data exists, it is difficult to access, due to unavailability of space science centres or institutions responsible for archiving this data. Results in chapter 6 showed that additional data improves prediction results during NN modelling. Physically, ionospheric parameters respond to changes in the solar activity which has an approximate sunspot cycle of 11 years. In order to fully assess a NN prediction model, different solar cycle stages should be taken into account and to achieve this, continuous archiving of GPS/GNSS data within both South Africa and beyond is required.

2. The GPS stations can be spaced within $\sim 3^\circ \times 3^\circ$ in latitude/longitude space to obtain simulated smoothed TEC maps. This was derived from the previous studies carried out by this author and others about TEC variability and predictability based on GPS-NN models (Habarulema et al., 2007, 2009a,b). Some other studies used grids of $1^\circ \times 1^\circ$ and $1^\circ \times 2^\circ$ in latitude and longitude for regional TEC mapping (Tulunay et al., 2006; Yilmaz et al., 2009). The advantage of regional TEC mapping and modelling using NNs is that once the model has been developed, the generation of TEC values over any region of interest can be performed at any resolution. The only disadvantage associated with this approach is the duration of the computation time which can be minimised by using high performance computing (HPC). It is therefore recommended that for the generation of high resolution smoothed TEC maps, HPC be explored in future.

8.3 Future work

Specifically in South Africa, GPS has provided an additional and relatively "cheap" method to characterise the ionosphere (e.g. Cilliers et al., 2004; Opperman et al., 2007; Moeketsi et al., 2007; McKinnell et al., 2007; Habarulema et al., 2007, 2010; McKinnell et al., 2010). It has been demonstrated that a comprehensive NN model can be developed to give insight into TEC variations at any point within South Africa. The current version of the SATECP model is

very useful in that it ensures complete datasets for applications involving investigation of TEC behaviour as a result of occurrence of unpredictable events such as magnetic storms (e.g. as shown in the previous chapter, figs. 7.10 and 7.11), especially in the absence of actual GPS data. Since not much work has been done concerning this study in the African region, future work would include the extension of the current work to cover the whole of Africa and produce predicted TEC maps of the continent, working towards improving the prediction accuracy and taking into account all solar activity levels for a fully inclusive prediction model. The simulated TEC maps of the continent will be among the space weather products provided by the Regional Warning Center for Africa. Currently, this is a method which the African region can use, since it has little or limited space weather infrastructure and archived ionospheric data. The method used in conjunction with extrapolation and interpolation techniques, allows for ionospheric behavioural changes to be derived and studied over a large geographical area (e.g. Komjathy, 1997; Yilmaz et al., 2009; Habarulema et al., 2010). It must be acknowledged that, from prior knowledge of ionospheric variability, the data used may not have covered the anticipated variations both geophysically and physically. The ionosphere owes its existence to the sun which has an approximately 11-year sunspot cycle. Throughout this study data used to develop the NN and linear regression models did not exceed a period of 8 years, which is obviously less than a sunspot cycle. This hindered the extensive validation of the SATECP model for both solar maximum and minimum outside the input space (with unseen data), a key factor to consider when assessing the generalisation capability of NNs in empirical modelling. The extension of this work as more data becomes available is therefore important in order to more accurately represent all geophysical parameters covering at least one solar cycle, and to improve predictions especially for storm conditions within the Southern Africa region. It is a well-established concept that data availability empowers the network to characterise the history of the physical parameter under consideration and this is a crucial requirement for NN applications (e.g. Haykin, 1994; Bishop, 1995).

References[1]

Araujo-Pradere, E. A. and Fuller-Rowell, T. J.: STORM: An empirical storm-time ionospheric correction model 2. Validation, Radio Science, 37(5), 1071, doi:10.1029/2002RS002 620, 2002.

Araujo-Pradere, E. A., Fuller-Rowell, T. J., and Codrescu, M. V.: STORM: An empirical storm-time ionospheric correction model, 1. Model description, Radio Science, 37(5), 1070, doi:10.1029/2001RS002 467, 2002.

Araujo-Pradere, E. A., Fuller-Rowell, T. J., and Bilitza, D.: Time Empirical Ionospheric Correction Model (STORM) response in IRI2000 and challenges for empirical modeling in the future, Radio Science, 39, RS1S24, doi:10.1029/2002RS002 805, 2004.

Barrile, V., Cacciola, M., and Cotroneo, F.: Multipath Reduction of GPS Measures through Heuristic Techniques of Compensation, in: Progress In Electromagnetic Research Symposium 2006, Cambridge, USA, pp. 528–532, 2006.

Baumjohann, W. and Treumann, R. A.: Basic Space Plasma Physics, Imperial College Press, 57 Shelton Street, Covert Garden London WC2H9HE, 1997.

Belehaki, A., Jakowski, N., and Reinisch, B. W.: Plasmaspheric electron content derived from GPS TEC and digisonde ionograms, Advances in Space Research, 33(6), 833–837, 2004.

Biktash, L. Z., Maruyama, T., and Nozaki, K.: The solar wind control of the equatorial ionosphere dynamics during geomagnetic storms, Advances in Space Research, 41(4), 562–568, 2008.

Bilitza, D.: International Reference Ionosphere 2000, Radio Science, 36(2), 261–275, 2001.

Bilitza, D.: International Reference Ionosphere 2000: examples of improvements and new futures, Advances in Space Research, 31(3), 757–767, 2003.

[1]The latex bibiliography style employed and adopted in this thesis was downloaded from the Copernicus Publications website (http://publications.copernicus.org).

Bilitza, D., Reinisch, B. W., Radicella, S. M., Pulinets, S., Gulyaeva, T., and Triskova, L.: Improvements of the International Reference Ionosphere model for topside electron density profile, Radio Science, 41, RS5S15, doi 10:1029/2005RS003 370, 2006.

Bishop, C. M.: Neural Networks for Pattern Recognition , Oxford University Press Inc.,New York, 1995.

Bodén, M.: A guide to recurrent neural networks and backpropagation, in: In the Dallas project, SICS Technical Report T2002:03, SICS, 2002.

Breed, A., Goodwin, G., Vandenberg, A.-M., Essex, E. A., Lynn, K. J. W., and Silby, J. H.: Ionospheric total electron content and slab thickness determined in Australia, Radio Science, 32(4), 1635–1643, 1997.

Cander, L. R.: Artificial neural network applications in ionospheric studies, Annali Di Geofisica, 5-6, 757–766, 1998.

Cander, L. R., Milosavljevic, M. M., Stankovic, S. S., and Tomasevic, S.: Ionospheric forecasting technique by artificial neural network, Electronic letters, 34(16), 1573–1574, 1998.

Carrano, C. and Groves, K.: Ionospheric Data Processing and Analysis, a tutorial at the workshop on Satellite Navigation Science and Technology for Africa, The Abdus Salam ICTP, Trieste, Italy, 2009.

Chan, A. H. Y. and Canon, P. S.: Nonlinear forecast of foF2: variation of model predictive accuracy over time, Annales Geophysicae, 20(7), 1031–1038, 2002.

Chen, F.: Introduction to Plasma Physics and Controlled Fusion, Plenum Press, 1984.

Chiu, M. C., Von-Mehlem, U. I., Willey, C. E., Betenbaugh, T. M., Maynard, J. J., Krein, J. A., Conde, R. F., Gray, W. T., Hunt, J. W., Mosher, L. E., McCullough, M. G., Panneton, P. E., Staiger, J. P., and Rodberg, E. H.: ACE spacecraft, Space Science Reviews, 86, 257–284, 1998.

Cilliers, P. J., Opperman, B. D. L., Mitchell, C. N., and Spencer, P. J.: Electron density profiles determined from tomographic reconstruction of total electron content obtained from GPS dual frequency data: first results from the South African network of dual frequency GPS receiver stations, Advances in Space Research, 34(9), 2049–2055, 2004.

Conway, A. J., Macpherson, K. P., Blacklaw, G., and Brown, J. C.: A neural network prediction of solar cycle 23, Journal of Geophysical Research, 103(A12), 29,733–29,742, 1998.

Davies, K.: Recent progress in satellite radio beacon studies with particular emphasis on the ATS-6 radio beacon experiment, Space Science Review, 25(4), 357–430, 1980.

Davies, K.: Ionospheric Radio, Peter Peregrinus Ltd., London, United Kingdom, 1990.

De Santis, A., de Franceschi, G., Zolesi, B., Pau, S., and Cander, L. R.: Regional mapping of the critical frequency of the F2 layer by spherical cap harmonic expansion, Annales Geophysicae, 9, 401–406, 1991.

Denton, M. H., Ulich, T., and Turunen, E.: Modification of midlatitude ionospheric parameters in the F2 layer by persistent high-speed solar wind streams, Space Weather, 7(S04006), doi:10.1029/2008SW000 443, 2009.

Elman, J. L.: Finding Structure in Time, Cognitive Science, 14, 179–211, 1990.

Fausett, L.: Fundamentals of Neural Networks; Architectures, Algorithms and Applications, Prentice-Hall, Inc. New Jersey 07632, 1994.

Gao, Y. and Liu, Z. Z.: Precise Ionospheric Modeling Using Regional GPS Network Data, Journal of Global Positioning Systems, 1(1), 18–24, 2002.

Girish, T. E., Jayachandran, B., and Shamsudeen, S. P.: Influence of solar wind on the TEC variations at mid and subauroral latitudes during sunspot maximum, Acta Geodaetica et Geophysica Hungarica, 32(3-4), 287–292, 1997.

Gledhill, J. A.: Thirty years of upper atmosphere research work at Rhodes, Part 1, The Radio Scientist, 2, 106–112, 1991.

Gledhill, J. A., Szendrei, M. E., and James, R. W.: The behaviour of the F-region of the ionosphere over Grahamstown during the partial solar eclipse of the 14th January 1945, Transactions of the Royal Society of South Africa, 31(3), 315–323, 1947.

GSFC: Goddard Space Flight Center, Space Physics Data Facility - Coordinated Data Analysis Web, `http://cdaweb.gsfc.nasa.gov/cdaweb/sp_phys/`; accessed in Feb, 2007.

GSFC: OMNI 2 dataset, `ftp://nssdcftp.gsfc.nasa.gov/spacecraft_data/omni/`; accessed in April, 2009.

Habarulema, J. B.: A feasibility study into Total Electron Content prediction using Neural Networks, Master's thesis, Rhodes University, Grahamstown, South Africa, 2007.

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 Terrestrial Physics, 69(15), 1842–1850, 2007.

Habarulema, J. B., McKinnell, L.-A., Cilliers, P. J., and Opperman, B. D. L.: Application of neural networks to South African GPS TEC modelling, Advances in Space Research, 43(11), 1711–1720, 2009a.

Habarulema, J. B., McKinnell, L.-A., and Opperman, B. D. L.: Towards a GPS-based TEC prediction model for Southern Africa with feed forward networks, Advances in Space Research, 44(1), 82–92, 2009b.

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(5), 2111–2125, 2009c.

Habarulema, J. B., McKinnell, L.-A., and Opperman, B. D. L.: TEC measurements and modelling over Southern Africa during magnetic storms ; a comparative analysis, Journal of Atmospheric and Solar Terrestrial Physics, 72(5-6), 509–520, 2010.

Haykin, S.: Neural Networks, A Comprehensive Foundation, Macmillan College Publishing Company, 1994.

Heilig, B., Lotz, S., Verő, J., Reda, J., Pajunpää, K., and Raita, T.: Empirically modelled Pc3 activity based on solar wind parameters, Annales Geophysicae, 28, 1703–1722, 2010.

Hernández-Pajares, M., Juan, J., and Sanz, J.: Neural network modelling of the ionospheric electron content at global scale using GPS, Radio Science, 32 (3), 1081–1090, 1997.

Hofmann-Wellenhof, B., Lichtenegger, H., and Collins, J.: Global Positioning System Theory and Practice , Springer-Verlag Wien New York, 1992.

Huang, C.-S. and Foster, J. C.: Variations of midlatitude ionospheric plasma density in response to an interplanetary schock, Geophysical Research Letters, 28(23), 4425–4428, 2001.

Huang, C.-S., Foster, J. C., Yumoto, K., Chau, J. L., and Veliz, O.: Prompt effects of solar wind variations on the inner magnetosphere and midlatitude ionosphere, Advances in Space Research, 36(12), 2407–2407, 2005.

Jakowski, N., Hocke, K., Schluter, S., and Heise, S.: Space weather effects detected by GPS based TEC monitoring, in: Workshop on Space Weather, WPP-155, ESTEC, Noordwijk,, pp. 241–244, 1999.

Joseyln, J. A.: Geomagnetic activity forecasting; the state of the art, Reviews of Geophysics, 33(3), 383–401, 1995.

King, J. and Papitashvilli, N.: OMNI 2 Data Documentation, NASA/NSSDC and QSS Group Inc., 2003.

Klobuchar, J.: Ionospheric Effects on GPS, Reprinted from GPS World, 1991.

Klobuchar, J. A.: Polarization of VHF Waves Emitted From Geostationary Satellites, Journal of Geophysical Research, 80 (31), 4387–4389, 1975.

Komjathy, A.: Global Ionospheric Total Electron Content Mapping Using the Global Positioning System, Ph.D. dissertation, Department of Geodesy and Geomatics Engineering, Technical Report No. 188, University of New Brunswick, Fredericton, New Brunswick, Canada, 1997.

Kouris, S. S., Fodadis, D. N., and Zolesi, B.: Specifications of the F-region variations for quiet and disturbed conditions, Physics and Chemistry of the Earth, 24(4), 321–327, 1999.

Kumar, V. V., Parkinson, M. L., Dyson, P. L., and Polglase, R.: Solar and geomagnetic activity on mid-latitude F-region electric fields, Annales Geophysicae, 26, 2911–2921, 2008.

Kutiev, I., Otsuka, Y., Saito, A., and Tsugawa, T.: Low-latitude total electron content enhancement at low geomagnetic activity observed over Japan, Journal of Geophysical Research, 112, A07 306, doi:10.1029/2007JA012 385, 2007.

Langley, R. B.: GPS, the Ionosphere, and the Solar Maximum, GPS World, pp. 44–49, 2000.

Leandro, R. F. and Santos, M. C.: A neural network approach for regional vertical total electron content modelling, Studia Geophysica et Geodaetica, 51(2), 279–292, 2007.

Liu, L., Wan, W., Lee, C. C., Ning, B., and Liu, J. Y.: The low latitude ionospheric effects of the April 2000 magnetic storm near the longitude 120°E, Earth, Planets Space, 56(6), 607–612, 2004.

Lyon, J. G.: The Solar Wind-Magnetosphere-Ionosphere System, Science Review, 288(5473), 1987–1991, 2000.

Macpherson, K. P., Conway, A. J., and Brown, J. C.: Prediction of solar and geomagnetic activity data using neural networks, Journal of Geophysical Research, 100(A11), 21,735–21,744, 1995.

Makela, J. J., González, S. A., MacPherson, B., Pi, X., Kelley, M. C., and Sultan, P. J.: Intercomparisons of total electron content measurements using Arecibo Incoherent Scatter Radar and GPS, Geophysical Research Letters, 27 (18), 2841–2844, 2000.

Mannucci, A. J., Wilson, B. D., and Yuan, D.-N.: An Improved ionospheric Correction Method for Wide-Area Augmentation Systems, in: ION GPS-95, Proceedings of the 8th International Technical Meeting of the Satellite Division of The Institute of Navigation Palm Springs, CA 12-15, September, The Institute of Navigation, Alexandria, Va., pp. 1199–1208, 1995.

Marra, S. and Marabito, F. C.: A New Technique for Solar Activity Forecasting using Recurrent Elman Networks, International Journal of Cognitive Intelligence, 3(1), 8–13, 2005.

Mathews, J. D.: Incoherent scatter radar observations of spread F producing ionospheric structures at arecibo, Journal of Atmospheric and Terrestrial Physics, 34 (6), 1119–1127, 1972.

McKinnell, L.-A.: A Neural Network based Ionospheric model for the bottomside electron density profile over Grahamstown, South Africa, PhD thesis of Rhodes University, Grahamstown, South Africa, 2002.

McKinnell, L.-A. and Poole, A. W. V.: A neural network based electron density model for the E layer, Advances in Space Research, 31(3), 589–595, 2003.

McKinnell, L.-A. and Poole, A. W. V.: Predicting the ionospheric F layer using neural networks, Journal of Geophysical Research, 109(A8), doi:10.1029/2004JA010 445, 2004.

McKinnell, L.-A., Opperman, B., and Cilliers, P. J.: GPS TEC and ionosonde TEC over Grahamstown, South Africa: First Comparisons, Advances in Space Research, 39(5), 816–820, 2007.

McKinnell, L.-A., Habarulema, J. B., Cilliers, P. J., and Opperman, B. D. L.: GPS assistance in modelling the Southern African ionosphere, in: Geoscience and Remote Sensing Symposium, 2009 IEEE International, IGARSS 2009, University of Cape Town, Cape Town, South Africa, pp. 883–885, 2010.

McNamara, L. F.: The Ionosphere, Communications, Surveillance, and Direction Finding, Krieger publishing company, 1991.

Meggs, R.: Mapping of ionospheric total electron content using global navigation satellite systems, Ph.D. thesis, University of Bath, Bath, UK, 2005.

Misra, P. and Enge, P.: Global Positioning System; Signals, Measurements and Performance, Ganga-Jumuna Press, P.O. Box 692 Lincoln, Massachusetts 01773, 2006.

Moeketsi, D.: Solar cycle effects on GNSS-derived ionospheric total electron content observed over Southern Africa, Ph.D. thesis, Rhodes University, Grahamstown, South Africa, 2008.

Moeketsi, D. M., Combrinck, W. L., McKinnell, L. A., and Fedrizzi, M.: Mapping GPS-derived ionospheric Total Electron Content over Southern Africa during different epochs of solar cycle 23, Advances in Space Research, 39(5), 821–829, 2007.

Mosert, M., Gende, M., Brunini, C., Ezquer, R., and Altadill, D.: Comparison of IRI TEC predictions with GPS and digisonde measurements at Ebro, Advances in Space Research, 39(5), 841–847, 2007.

Müller, B. and Reinhardt, J.: Neural Networks; An introduction, ISBN 3-540-97614-0 Springler-Verlag Berlin Heidelberg New York, Springer-Verlag Berlin Heidelberg, 1990.

Ondoh, T. and Marubashi, K.: Wave summit course; Science of Space Environment, Ohmsha, Ltd. 3-1 Kanda Nishiki-cho, Chiyoda-ku, Tokyo 101-8460 Japan, 2000.

Opperman, B.: Reconstructing Ionospheric TEC over South Africa using signals from a Regional GPS network, PhD thesis, Rhodes University, Grahamstown, South Africa, 2008.

Opperman, B. D. L., Cilliers, P. J., McKinnell, L.-A., and Haggard, R.: Development of a Regional GPS-based Ionospheric TEC model for South Africa, Advances in Space Research, 39(5), 808–815, 2007.

Oyeyemi, E. O., McKinnell, L.-A., and Poole, A. W. V.: Near-real time foF2 predictions using neural networks, Journal of Atmospheric and Solar-Terrestrial Physics, 68(16), 1807–1818, 2006.

Poole, A. W. V. and Evans, G. P.: Advanced sounding: (ii) First results from an advanced chirp ionosonde, Radio Science, 20, 1617–1623, 1985.

Poole, A. W. V. and McKinnell, L.-A.: On the predictability of foF2 using neural networks, Radio Science, 35, 225–234, 2000.

Poole, A. W. V. and Poole, M.: Long-term trends in foF2 over Grahamstown using Neural Networks, Annals of Geophysics, 45, 155–161, 2002.

Reczko, M., Riedmiller, M., Seemann, M., Ritt, M., DeCoster, J., Biedermann, J., Danz, J., Wehrfritz, C., Werner, R., Berthold, M., and Orsier, B.: External contributions to Stuttgart Neural Network Simulator (SNNS), User Manual, Version 4.2, Universities of Stuttgart and Tübingen, Germany, and the European Particle Research Lab, CERN, Geneva, Switzerland, 1998.

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

Reinisch, B. W., Galkin, I. A., Khmyrov, G., Kozlov, A., and Kitrosser, D. F.: Automated collection and dissemination of ionospheric data from the digisonde network, Advances in Radio Science, 2, 241–247, 2004.

Rishbeth, H., Kohl, H., and Barclay, W.: A history of ionospheric physics and radio communications, in: Modern Ionospheric Science, pp. 4–31, 1996.

Sandwell, D. T.: Biharmonic spline interpolation of GEOS-3 and SEASAT altimeter data, Geophysical Research Letters, 14(2), 139–142, 1987.

Sarma, A. D. and Madhu, T.: Modelling of foF2 using neural networks at an equatorial anomaly station, Current Science, 89(7), 1245–1247, 2005.

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.

Senalp, E. T., Tulunay, E., and Tulunay, Y.: Total electron content (TEC) forecasting by Cascade Modeling, A possible alternative to the IRI-2001, Radio Science, 43, RS4016, doi:10.1029/2007RS003 719, 2008.

Sethia, G., Deshpande, M. R., and Rastogi, R. G.: The solar wind influences plasmasphere electron content, Nature, 276, 482, 1978.

Shue, J.-H., Song, P., Russell, C. T., Steinberg, J. T., Chao, J. K., Zastenker, G., Vaisberg, O. L., Kokubun, S., Singer, H. J., Detman, T. R., and Kawano, H.: Magnetopause location under extreme solar wind conditions, Journal of Geophysical Research, 103(A8), 17,691–17,700, 1998.

Sibanda, P.: Challenges in Topside ionospheric modelling over South Africa, Ph.D. thesis, Rhodes University, Grahamstown, South Africa, 2010.

Stankov, S. M., Kutiev, I., Jakowski, N., and Wehrenpfennig, A.: A new method for total electron content forecasting using Global Positioning System measurements, in: Proc. ESA Space Weather Workshop, Noordwijk, The Netherlands, pp. 169–172, 2001.

Stone, E. C., Frandsen, A. M., Mewaldt, R. A., Christian, E. R., Margolies, D., Ormes, J. F., and Snow, F.: The Advanced Composition Explorer, Space Science Reviews, 86 (1-4), 1–22, 1998.

Tsurutani, B. T. and Gonzalez, W. D.: The interplanetary causes of magnetic storms: A review, in: Magnetic Storms, edited by: Tsurutani, B. T., Gonzalez, W. D., Kamide, Y., and Arballo, J. K.; Geophysical Monograph 98, 1997.

Tulunay, E., Senalp, E. T., Radicella, S. M., and Tulunay, Y.: Forecasting total electron content maps by neural network technique, Radio Science, 41, doi:10.1029/2005RS003 285, 2006.

Vandegriff, J., Wagstaff, K., Ho, G., and Plauger, J.: Forecasting space weather: Predicting interplanetary shocks using neural networks, Advances in Space Research, 36, 2323–2327, 2005.

Verő, J.: Geomagnetic pulsations and parameters of the interplanetary medium, Journal of Atmospheric and Terrestrial Physics, 42(4), 371–380, 1980.

Watson, M.: Common LISP Modules; Artificial Intelligence in the Era of Neural networks and Chaos Theory, ISBN 3-540-97614-0 Springler-Verlag Berlin Heidelberg New York, Springer-Verlag New York, Inc., 1991.

Weigel, R. S., Vassiliadis, D., and Klimas, A. J.: Coupling of the solar wind to temporal fluctuations in ground magnetic fields, Geophysical Research Letters, 29(19), 1945, doi:10.1029/2002GL014 740, 2002.

Weigel, R. S., Klimas, A. J., and Vassiliadis, D.: Solar wind coupling to and predictability of ground magnetic fields and their time derivatives, Journal of Geophysical Research, 108(A7), 1298, doi:10.1029/2002JA009 627, 2003.

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

Wu, J.-G. and Lundstedt, H.: Neural network modeling of solar wind-magnetosphere interaction, Journal of Geophysical Research, 102(A7), 14 457–14 466, 1997.

Yilmaz, A., Akdogan, K. E., and Gurun, M.: Regional TEC mapping using neural networks, Radio Science, 44, RS3007, doi:10.1029/2008RS004 049, 2009.

Yizengaw, E., Moldwin, M. B., Dyson, P. L., and Immel, T. J.: Southern Hemisphere ionosphere and plasmasphere response to the interplanetary shock event of 29-31 October 2003, Journal of Geophysical Research, 110, A09S30, doi:10.1029/2004JA010920, 2005.

Zell, A., Mamier, G. M., Vogt, M., Mache, N., Hübner, R., Döring, S., Herrmann, K.-U., Soyez, T., Schmalzl, M., Sommer, T., Hatzigeorgiou, A., Posselt, D., Schreiner, T., Kett, B., Clemente, G., Wieland, J., and Gatter, J.: Stuttgart Neural Network Simulator (SNNS), User Manual, Version 4.2, Universities of Stuttgart and Tübingen, Germany, and the European Particle Research Lab, CERN, Geneva, Switzerland, 1998.