



**MODELLING OF GNSS DERIVED VERTICAL
TOTAL ELECTRON CONTENT VIA ARTIFICIAL
NEURAL NETWORKS: A CASE STUDY
IN BRAZIL**

ARTHUR AMARAL FERREIRA

**DISSERTAÇÃO DE MESTRADO EM ENGENHARIA DE SISTEMAS
ELETRÔNICOS E DE AUTOMAÇÃO
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

FACULDADE DE TECNOLOGIA

UNIVERSIDADE DE BRASÍLIA

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

**MODELLING OF GNSS DERIVED VERTICAL
TOTAL ELECTRON CONTENT VIA ARTIFICIAL
NEURAL NETWORKS: A CASE STUDY
IN BRAZIL**

ARTHUR AMARAL FERREIRA

Orientador: PROF. DR. RENATO ALVES BORGES, ENE/UNB

**DISSERTAÇÃO DE MESTRADO EM ENGENHARIA DE SISTEMAS
ELETRÔNICOS E DE AUTOMAÇÃO**

**PUBLICAÇÃO PPGEA.DM - 687/2018
BRASÍLIA-DF, 07 DE FEVEREIRO DE 2018.**

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

MODELLING OF GNSS DERIVED VERTICAL TOTAL ELECTRON
CONTENT VIA ARTIFICIAL NEURAL NETWORKS: A CASE STUDY
IN BRAZIL

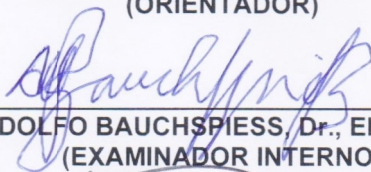
ARTHUR AMARAL FERREIRA

DISSERTAÇÃO DE MESTRADO SUBMETIDA AO DEPARTAMENTO DE ENGENHARIA
ELÉTRICA DA FACULDADE DE TECNOLOGIA DA UNIVERSIDADE DE BRASÍLIA, COMO
PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE.

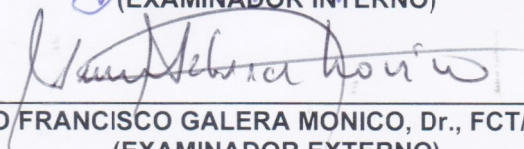
APROVADA POR:



RENATO ALVES BORGES, Dr., ENE/UNB
(ORIENTADOR)



ADOLFO BAUCHSPIESS, Dr., ENE/UNB
(EXAMINADOR INTERNO)



JOÃO FRANCISCO GALERA MÔNICO, Dr., FCT/UNESP
(EXAMINADOR EXTERNO)

Brasília, 07 de fevereiro de 2018.

FICHA CATALOGRÁFICA

ARTHUR AMARAL FERREIRA

MODELLING OF GNSS DERIVED VERTICAL TOTAL ELECTRON CONTENT VIA ARTIFICIAL NEURAL NETWORKS: A CASE STUDY IN BRAZIL

2018xii, 73p., 201x297 mm

(ENE/FT/UnB, Mestre, Engenharia de Sistemas Eletrônicos e de Automação, 2018)

Dissertação de Mestrado - Universidade de Brasília

Faculdade de Tecnologia - Departamento de Engenharia Elétrica

- | | |
|-------------------------------|--|
| 1. Artificial Neural Networks | 3. Ionosphere |
| 2. Total Electron Content | 4. Global Navigation Satellite Systems |
-

REFERÊNCIA BIBLIOGRÁFICA

ARTHUR AMARAL FERREIRA (2018) MODELLING OF GNSS DERIVED VERTICAL TOTAL ELECTRON CONTENT VIA ARTIFICIAL NEURAL NETWORKS: A CASE STUDY IN BRAZIL. Dissertação de Mestrado em Engenharia de Sistemas Eletrônicos e de Automação, Publicação 687/2018, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 73p.

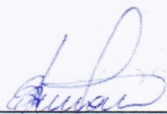
CESSÃO DE DIREITOS

AUTOR: ARTHUR AMARAL FERREIRA

TÍTULO: MODELLING OF GNSS DERIVED VERTICAL TOTAL ELECTRON CONTENT VIA ARTIFICIAL NEURAL NETWORKS: A CASE STUDY IN BRAZIL.

GRAU: Mestre ANO: 2018

É concedida à Universidade de Brasília permissão para reproduzir cópias desta dissertação de Mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor se reserva a outros direitos de publicação e nenhuma parte desta dissertação de Mestrado pode ser reproduzida sem a autorização por escrito do autor.



ARTHUR AMARAL FERREIRA

Conjunto 5, QRS 214, Setor Militar Urbano, Brasília-DF, Brasil.

Acknowledgments

First, I would like to thank God for giving me health, wisdom, strength and the family I have.

I thank my parents, Gileno and Edinalva, for inspiring and supporting me during my entire life. Their unconditional love gives me the motivation to keep pursuing my dreams. I thank my brother Douglas and my sister Thaís that provide me with joy and happiness during the moments of trouble. I also would like to thank Aline for the encouragement and support during my academic journey.

I would like also to thank my supervisor, Prof. Renato Borges, for his advice, teaching, dedication and for being solicitous and patient whenever I needed help. His guidance was very important for the accomplishment of this work. I thank Dr. Claudia Paparini who have been very solicitous during the last two years, contributing to the development of this work.

I also would like to thank the International Centre for Theoretical Physics (ICTP) for the opportunity I had to attend the training workshops in 2016 and 2017. These activities played a very important role in the development of this work. I thank Dr. Sandro M. Radicella and Dr. Luigi Ciruolo for the knowledge transmitted during our meetings and for their contributions to the studies performed during this period.

I thank the Russian joint stock company JC«RPC«PSI» for provide the access to the GLONASS data obtained from the GLONASS R& D network in Brazil. Also, I would like to thank the Brazilian Institute of Geography and Statistics (IBGE) for providing the data from the RBMC network.

I would like to thank the Coordination for the Improvement of Higher Education Personnel (CAPES) for the financial support that enabled me to dedicate exclusively to the development of this research.

Arthur Amaral Ferreira

Abstract

One of the main error sources on Global Navigation Satellite Systems (GNSS) positioning solutions for users of single frequency receivers is the propagation refraction of the GNSS signals as they pass through the ionosphere. The estimation of the Total Electron Content (TEC) is very important for the correction of ionosphere propagation effects on GNSS signals. In order to correct the ionospheric range errors, GNSS single-frequency users need to rely on TEC models. In this framework, the present investigates the use of Artificial Neural Network models (ANN) to estimate TEC derived from GNSS measurements in Brazil. More specific, the investigations start the development of a regional model that can be used to determine the vertical TEC ($vTEC$) over Northeast, Central-West and South regions of Brazil, aiming future applications on a near real-time frame estimations and short-term forecasting.

This work uses GNSS data from the GLONASS network for research and development, and from the Brazilian Network for Continuous Monitoring of the GNSS (RBMC). The input parameters of the ANN models are based on features known to influence TEC values, including the geographic location of the GNSS receiver, geomagnetic activity, seasonal and diurnal variations, and solar activity. The proposed ANN model is used to estimate the GNSS TEC values at void locations, where no dual-frequency GNSS receiver that may be used as a source of data for GNSS TEC estimation is available.

Different analyses are carried out divided into three case studies. These analyses include spatial performance evaluation, evaluation of different ANN structures, short-term forecasting ability and performance comparison against CODE (Center for Orbit Determination in Europe) Global Ionospheric Maps during the geomagnetic storm registered on 13th and 14th October 2016. The results obtained from the described analysis suggest that the proposed ANN models provides good spatial performance and presents to be a promising tool for short-term forecasting applications.

Resumo

Uma das principais fontes de erro no posicionamento baseado em Sistemas Globais de Navegação por Satélite (GNSS) para usuários de receptores de uma frequência é o atraso de propagação nos sinais GNSS ao atravessarem a ionosfera. Esse atraso, em uma aproximação de primeira ordem, é diretamente proporcional ao Conteúdo Total de Elétrons (TEC). Assim, estimar o TEC é uma tarefa bastante relevante para correção dos efeitos ionosféricos sobre a propagação dos sinais. Para corrigir os erros de distância devido à ionosfera, os usuários de receptores GNSS de uma única frequência necessitam de modelos que representem o TEC. Neste cenário, este trabalho propõe a utilização de Redes Neurais Artificiais (ANN) para estimar o TEC obtido a partir de medidas GNSS na região do Brasil. As investigações apresentadas neste trabalho iniciam o desenvolvimento de um modelo regional que possa ser usado para determinar o TEC vertical sobre as regiões Nordeste, Centro-Oeste e Sul do Brasil, visando futuras aplicações em estimação próxima a tempo real e em previsão de curto prazo.

Neste trabalho são utilizados dados GNSS das redes GLONASS para pesquisa e desenvolvimento, e da Rede Brasileira de Monitoramento Contínuo dos Sistemas GNSS (RBMC). Os parâmetros de entrada da rede neural baseiam-se em fatores que influenciam os valores do TEC, incluindo localização geográfica do receptor GNSS, atividade geomagnética, variações sazonais e diurnas e atividade solar. O modelo de ANN proposto é utilizado para estimar os valores de GNSS TEC vertical em regiões desprovidas de receptores GNSS de duas bandas de frequência que possam ser utilizados para tal fim.

Diferentes análises são realizadas, divididas em três estudos de caso. Estas análises incluem a avaliação de desempenho espacial, avaliação de diferentes estruturas ANN, habilidade de previsão em curto-prazo e comparação de desempenho em relação aos Mapas Ionosféricos Globais (*Global Ionospheric Maps*) fornecidos pelo Centro para Determinação de órbita na Europa (CODE) durante a tempestade geomagnética registrada nos dias 13 e 14 de Outubro de 2016. Os resultados obtidos a partir das análises conduzidas sugerem que os modelos de NN propostos fornecem bom desempenho espacial e apresentam-se como ferramentas promissoras para aplicações de previsão de TEC de curto-prazo.

LIST OF CONTENTS

1	INTRODUCTION	1
1.1	GOALS AND CONTRIBUTIONS	2
1.2	PRESENTATION OF THE MANUSCRIPT	3
2	FUNDAMENTALS	4
2.1	GLOBAL NAVIGATION SATELLITE SYSTEMS	4
2.1.1	GLOBAL POSITIONING SYSTEM (NAVSTAR-GPS)	5
2.1.2	GLONASS	7
2.1.3	GALILEO	8
2.1.4	BEIDOU	8
2.2	THE GNSS OBSERVABLES	9
2.2.1	THE PSEUDORANGE	9
2.2.2	THE CARRIER PHASE	12
2.3	GNSS OBSERVABLES ERROR SOURCES	12
2.3.1	CLOCK ERRORS	13
2.3.2	MULTIPATH	14
2.3.3	TROPOSPHERIC EFFECT	14
2.3.4	IONOSPHERIC EFFECT	16
2.4	ARTIFICIAL NEURAL NETWORKS	22
2.4.1	ROSENBLATT PERCEPTRON	23
2.4.2	MULTILAYER PERCEPTRON	23
3	RESULTS AND ANALYSIS	27
3.1	PRELIMINARIES	27
3.1.1	NN MODEL INPUT PARAMETERS	27
3.1.2	CALIBRATION TECHNIQUES	29
3.1.3	THE ANN MODELS	30
3.1.4	NN PERFORMANCE EVALUATION	31
3.2	CASE STUDY 1	32
3.2.1	TRAINING THE NN	33
3.2.2	RESULTS AND DISCUSSION	34
3.3	CASE STUDY 2	35
3.3.1	TRAINING THE NN	37

3.3.2	RESULTS AND DISCUSSIONS.....	37
3.4	CASE STUDY 3	38
3.4.1	TRAINING THE NN	39
3.4.2	RESULTS AND DISCUSSION	40
4	CONCLUSION	49
	REFERENCES	51
	APPENDICES	56
A	LIST OF PUBLICATIONS	57
A.1	DIRECTLY RELATED PUBLICATIONS.....	57
A.2	INDIRECTLY RELATED PUBLICATIONS	57
B	LINEARIZATION OF THE OBSERVABLE EQUATION.....	58
C	ADDITIONAL TROPOSPHERIC EFFECT INFORMATION	60
C.1	TROPOSPHERIC ATTENUATION AS A FUNCTION OF THE ELEVATION ANGLE	60
C.2	OTHER TROPOSPHERIC EFFECTS	61
C.3	EMPIRICAL MODELS AND MAPPING FUNCTIONS	62
D	ADDITIONAL IONOSPHERIC EFFECT INFORMATION.....	66
E	OTHER SOURCES OF ERRORS	68
F	POSITIONING TECHNIQUES	71

LIST OF FIGURES

2.1	Receiver positioning calculation.	9
2.2	Use of replica code to determine the satellite code travelling time (MONICO, 2008; KAPLAN; HEGARTY, 2006) (<i>adapted</i>).....	10
2.3	Multipath illustration.	14
2.4	Example of actual monthly average time-delay data, along with cosine model fit to data (KLOBUCHAR, 1987).	21
2.5	Artificial neuron model (HAYKIN, 2009) (<i>adapted</i>).	22
2.6	Hyperplane as a decision boundary for a two-dimensional, two-class pattern-classification problem (HAYKIN, 2009) (<i>adapted</i>).	24
3.1	Schematic representation of one ANN structure.	31
3.2	RMSE values for each NN structure tested - Case study 1.	34
3.3	Calibrated vTEC and NN vTEC on the September equinox of 2016 for the South region - Case study 1. (a) SMAR station, (b) POAL station.....	35
3.4	Calibrated vTEC and NN vTEC on the September equinox of 2016 for the Central-West region - Case study 1. (a) GOGY station. (b) MTNX station.....	36
3.5	Positions in the world map of the stations under investigation - Case study 2. Green and blue dots represent training and testing station, respectively.	37
3.6	Performance of the ANN during the training and validation procedures - Case study 2.	38
3.7	Calibrated vTEC and NN vTEC on GOGY station - Case Study 2.	38
3.8	Positions in the world map of the stations under investigation - Case study 3. Green and blue dots represent training and testing stations, respectively.	40
3.9	Calibrated vTEC and NN vTEC on the PEAFF station - Case Study 3. (a) Period A. (b) Period B.....	41
3.10	Calibrated vTEC and NN vTEC on GOGY station - Case study 3. (a) Period A. (b) Period B.	43
3.11	Calibrated vTEC and NN vTEC on SMBJ station - Case study 3. (a) Period A. (b) Period B.	44
3.12	Calibrated vTEC and NN vTEC for short-term forecasting of doy 292 - Case study 3. (a) PEAFF station. (b) GOGY station. (c) SMBJ station.	45

3.13	Calibrated vTEC, NN vTEC, and GIM derived vTEC from doy 286 to 288 on year 2016 - Case study 3. (a) PEAUF station. (b) GOGY station. (c) SMBJ station.	47
C.1	Path length L through an uniform shell troposphere at elevation angle ϵ (SPILKER, 1996) (<i>adapted</i>).	60
C.2	(Tropospheric attenuation versus elevation angle (SPILKER, 1996) (<i>adapted</i>)).	61

LIST OF TABLES

2.1	Carrier frequency per signal in the GPS (MONICO, 2008).	6
2.2	Carrier frequency per signal in the Galileo system (The European Commission, 2016).....	8
2.3	Signal characteristics in the BeiDou system (Novatel Inc., 2015).	9
2.4	Characteristics of the different regions of the ionosphere (KLOBUCHAR, 1996).....	17
3.1	NN structures under investigation - Case study 1.....	33
3.2	Summarized information of the stations under investigation in the South region - Case study 1.	33
3.3	Summarized information of the stations under investigation in the Central-West region - Case study 1.	35
3.4	Spatial performance of the NN model - Case Study 1.....	36
3.5	Summarized information of the stations under investigation - Case study 2.	36
3.6	Spatial performance of the NN model - Case study 2.	38
3.7	Summarized information of the stations under investigation in the Northeast region - Case study 3.	40
3.8	Summarized information of the stations under investigation in the Central-West region - Case study 3.	42
3.9	Summarized information of the stations under investigation in the South region - Case study 3.....	43
3.10	Spatial performance of the NN model - Case study 3.	45
3.11	Short-term forecasting performance of the NN model - Case study 3.....	46
3.12	Registered Kp-index values.	46
3.13	Performance evaluation of NN vTEC and GIM vTEC with respect to calibrated TEC from doy 286 to 288.	48
D.1	Classification of the scintillation levels.	67
E.1	Uncertainty of types of ephemeris (HOFMANN-WELLENHOF; LICHTENEGGER; WASLE, 2008) (<i>adapted</i>).....	69

Acronyms

ANN	Artificial Neural Network
CDDIS	Crustal Dynamics Information Systems
CODE	Center for Orbit Determination in Europe
CW	Central-West region
DCB	Differential Code Bias
DoD	U.S. Department of Defense
doy	Day of the year
Dst	Disturbance storm time index
EMBRACE	Brazilian Study and Monitoring of the Space Weather
ESA	European Space Agency
GIM	Global Ionospheric Map
GLONASS	Global'naya Navigatsionnaya Sputnikovaya Sistema
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GPT2	Global Pressure and Temperature Model 2
GPT2w	Global Pressure and Temperature Model 2 wet
ICTP	International Centre for Theoretical Physics
IFB	Interfrequency bias
IONEX	IONosphere map Exchange format
ITG	Improved Tropospheric Grid
JPL	Jet Propulsion Laboratory
LORAN	LONg-RANge navigation system
MLP	Multilayer Perceptron
MSE	Mean Square Error
NE	Northeast region
NOAA	National Oceanic and Atmospheric Administration
NRMSE	Normalized Root Mean Square Error
RBMC	Brazilian Network for Continuous Monitoring of GNSS (RBMC)
RMSE	Root Mean Square Error
SO	South region

sTEC	Slant Total Electron Content
TEC	Total Electron Content
TECU	Total Electron Content Unit
UNB3	University of New Brunswick tropospheric model 3
UNB3m	University of New Brunswick tropospheric model 3m
UNB4	University of New Brunswick tropospheric model 4
UPC	Polytechnic University of Catalonia
RIM	Regional Ionospheric Map
RTCA-MOPS	Radio Technical Commission for Aeronautics - Minimum Operational Performance Standards
vTEC	Vertical Total Electron Content
WDC	World Data Center for Geomagnetism

List of Symbols

c	Speed of light in vacuum
sv	Satellite vehicle
r	Receiver
f	Signal frequency
L_1	GPS carrier at 1574.42 MHz
L_2	GPS carrier at 1227.60 MHz
L_5	GPS carrier at 1176.45 MHz
$E1$	Galileo carrier at 1575.42 MHz
$E6$	Galileo carrier at 1278.75 MHz
$E5_a$	Galileo carrier at 1176.45 MHz
$E5_b$	Galileo carrier at 1207.14 MHz
$B1$	Galileo carrier at 1207.14 MHz
$B2$	Galileo carrier at 1278.75 MHz
$B3$	Galileo carrier at 1268.52 MHz
X_r	X-coordinate of the receiver position in the ECEF coordinate system
Y_r	Y-coordinate of the receiver position in the ECEF coordinate system
Z_r	Z-coordinate of the receiver position in the ECEF coordinate system
X^{sv}	X-coordinate of the satellite position in the ECEF coordinate system
Y^{sv}	Y-coordinate of the satellite position in the ECEF coordinate system
Z^{sv}	Z-coordinate of the satellite position in the ECEF coordinate system
G^{st}	Satellite generated code
$G_r t$	Receiver generated replica code
t^{sv}	Satellite time system
t_r	Receiver time system
dt^{sv}	Satellite clock error
dt_r	Receiver clock error
P_r^{sv}	Pseudorange between the satellite sv and the receiver r
ε_P	Pseudorange measurement error
ρ_r^{sv}	Geometric distance between the satellite and the receiver r
T_r^{sv}	Tropospheric delay
I_r^{sv}	Ionospheric delay
dm_r^{sv}	Multipath effect

$\Phi_{r,sv}$	Carrier phase observable
Δ	Variation
doy	Day of year
λ	Wavelength of the carrier
N_{sv}	Ambiguity
$\varepsilon_{\Phi_r^{sv}}$	Carrier phase measurement error
a_0	satellite clock coefficient
a_1	satellite clock drift coefficient
a_2	satellite clock frequency drift coefficient
Δ_R	Correction for the satellite clock relativistic effect
ϵ	Elevation angle
N_T	Tropospheric refractivity
N_H	Hydrostatic Tropospheric refractivity component
N_W	Wet Tropospheric refractivity component
N_e	Electron density
T_{ZH}	Zenith hydrostatic tropospheric delay
T_{ZW}	Zenith wet tropospheric delay
n_{gr}	Group refractive index
d_{Igr}	Ionospheric group delay
d_I	Carrier phase advance
n	Phase refractive index
an_0	Effective Ionization Level 1st order parameter - Nequick model
an_1	Effective Ionization Level 2st order parameter - Nequick model
an_2	Effective Ionization Level 3rd order parameter - Nequick model
A_z	Effective Ionization Level - Nequick model
μ	Modified dip (modip)
x_m	m -esim neuron input
w	Synaptic weight
b	Bias of a neuron
v	Induced local field of a neuron
d_i	Desired output of a neuron
η	Learning rate
Φ	Activation function
y	Output of a neuron
e	Output error of a neuron
\mathcal{T}	Set of training samples
dn/df	First derivative of the phase refractive index with respect to the frequency
$mh(\epsilon)$	Mapping function to convert hydrostatic zenith tropospheric delay to slant hydrostatic tropospheric delay
$mw(\epsilon)$	Mapping function to convert wet zenith tropospheric delay to slant wet tropospheric delay
α	Average error

ϵ	Average relative error
$F_{10.7}$	Solar radio flux
$NNvTEC$	vertical TEC estimated by the Neural Network
a	Slope parameter of the logistic sigmoid function
exp	Exponential function

Chapter 1

Introduction

With the development of maritime navigation, the need of knowledge of the vessels position on the Earth became crucial. The exploration and the conquest of new territories in a way that the vessels movement were safe required the abilities to move from one place to another and also to determine geographic positions (MONICO, 2008). Several developments have been made to allow the navigation activities including the invention of some tools such as the compass, quadrant, astrolabe and also the development of radio navigation systems such as the LORAN (LONg-RANge navigation system) and the OMEGA. In this scenario, what has most significantly changed navigation techniques is the advent of Global Navigation Satellite Systems (GNSS), which started with the launch of the U.S. Department of Defense Global Positioning System (GPS) in the late 1970s (MONICO, 2008; Novatel Inc., 2015).

Following the development of the GPS, other GNSS have been made available, such as the GLONASS, Galileo and Beidou making the use of this technology widespread, including applications in navigation, air-craft landing, high-precision agriculture and others. Nowadays, vehicles, whether on land, in the air or at sea, routinely rely on the accurate positioning information provided by GNSS technology. In fact, the ready adoption of the technology, from mining to unmanned, and the increasingly complex requirements for positioning, anywhere and anytime, are driving innovation in the industry that includes the integration of GNSS technology with a variety of other sensors and methodologies (Novatel Inc., 2015).

The observables used to determine position using the GNSS technology are subject to different errors, that can be classified into random, systematic and gross errors. The sources of these several errors can be related to the satellite (e.g orbit errors, satellite clock errors), the receiver/antenna (e.g receiver clock errors), station (e.g. coordinates errors, multipath) and errors related to the signal propagation (MONICO, 2008).

When traveling through a static electric or magnetic fields in a linear medium such as a vacuum an electromagnetic wave is not affected. However, when traveling in a dispersive medium, such as the atmosphere, different aspects cause variation in the propagation speed,

polarization and signal power (BORRE; STRANG, 2012; HOQUE; JAKOWSKI, 2015). The medium where the GNSS signals propagate consists of the troposphere and ionosphere, essentially. Each one of these atmospheric layers has its particular characteristics and has particular impact on the propagation of the GNSS signals, leading to errors in the positioning determination (MONICO, 2008).

In the context of GNSS L band signals, the ionospheric refraction introduces most of the delay that may cause range errors in the positioning system of up to 100 m (JAKOWSKI et al., 2011). The first order ionospheric delay is directly proportional to the Total Electron Content (TEC) that is the number of electrons in a column with cross-sectional area of 1 m^2 along the path from the satellite to the receiver (CESARONI et al., 2015). The ionospheric effect in GNSS applications is even worst in the equatorial and low-latitude regions, since in these areas the TEC presents strong temporal and spatial variation due to mainly three different dynamic processes: the equatorial ionization anomaly, post-sunset plasma enhancement and evening plasma bubbles (TAKAHASHI et al., 2014).

Some benefits of knowing the correct TEC value within a good spatial resolution is related with the improvement of accuracy in global navigation satellite systems positioning solutions, as well as a better understanding of the different parameters that affect it, such as solar and magnetic activities, and the ability for monitoring and forecast space weather events (DENARDINI; DASSOB; G.-ESPARZAD, 2016a; DENARDINI; DASSOB; G.-ESPARZAD, 2016b). In this context, the use of ANNs has provided good results in applications for regional TEC modelling being capable of recovering TEC values with good performance, higher than 80% on average (LEANDRO; SANTOS, 2007; HABARULEMA et al., 2009; MACHADO, 2012). This fact is related to the abilities of an ANN to learn, generalize and adapt to different patterns of input/output sets with nonlinear behavior (HAYKIN, 1999).

Some works have been carried out by using neural networks for prediction and modelling of ionospheric parameters. An ANN model for Brazil, considering only the geographic position as ANN input is proposed in LEANDRO; SANTOS, (2007). TEC prediction and modelling of TEC in South Africa and Nigeria, using ANN, can be found in HABARULEMA; MCKINNELL; CILLIERS, (2007), HABARULEMA et al., (2009) and OKOH et al., (2016), respectively. MACHADO, (2012) presents a methodology to predict the vertical TEC, regionally, by using an ANN structure, aiming to generate virtual reference stations, to be employed in positioning techniques, over São Paulo state, in Brazil.

1.1 Goals and contributions

In this framework, this work aims to use ANNs to estimate TEC values based on GNSS measurements in three Brazilian sectors. The idea is based on the work of Leandro and Santos 2007, but considering different activation functions for the ANNs, more input parameters

similar to the approach presented in Habarulema et al. 2009 and using data from the new GLONASS R&D network. In order to assess the performance of the proposed NN model, the investigation is divided into three studies of case, each one presenting different approaches and characteristics.

Different to the approach adopted by LEANDRO; SANTOS, (2007), the present work considers input parameters related to time variability of the ionospheric activity. Since TEC is influenced by the solar activity, diurnal and seasonal variations, magnetic field of the Earth and geographic location of the GNSS receiver (HOFMANN-WELLENHOF; LICHTENEGGER; WASLE, 2008), the input parameters are chosen to include this information. Furthermore, the spatial performance of the proposed ANN model is assessed during the geomagnetic storm registered on October 13th and 14th 2016 and compared with Global Ionospheric Maps, provided by the Center for Orbit Determination in Europe (CODE). In addition, the ANN model ability to perform short-term forecasting using low amount of data is assessed.

It is worth mentioning that this work is the first one using data from the GLONASS R&D network recently inaugurated in Brazil. This network consists of three ground stations (BRAJ, RSFJ, SMBJ) inaugurated from the middle of 2014 to the beginning of 2016. These stations are installed in different regions of Brazil (North-east, Center-West and South) allowing to investigate the applicability of the ANN model in different latitudes.

1.2 Presentation of the manuscript

Chapter 2 presents an overview to provide some fundamental information about the GNSS technology, including some features of the global navigation satellite systems available, some positioning calculation concepts and also some sources of errors that affect the GNSS positioning solutions. A brief discussion about Artificial Neural Networks is also presented in the chapter, emphasizing the Multilayer Perceptron (MLP) structure that formed the basis for the v TEC estimations performed in this work.

Chapter 3 presents the three case of studies conducted in this work, giving particular information about each case. The results obtained by using the ANN model are presented and discussed. This manuscript is concluded in Chapter 4 summarizing the obtained results and suggesting ideas for future work.

Chapter 2

Fundamentals

The purpose of this chapter is to present a brief overview of the global navigation satellite systems available to the user, some characteristics of each system and also the basic GNSS positioning concepts. Following this brief discussion, the sources of errors that affect the GNSS observables are presented. More emphasis is given to the ionospheric refraction which is one of the major sources of error on GNSS and is directly proportional to the TEC, the parameter to be estimated in this work.

A brief discussion about the Artificial Neural Networks is presented, more specific the Multilayer Perceptron (MLP) structure that formed the basis for the TEC estimations carried out in this manuscript.

2.1 Global Navigation Satellite Systems

The mankind has always been interested in positioning determination. Human beings have always interested to know their locations in space; at the beginning they are interested to know their positioning in the vicinity of home. Since then the area of interest has increased to all the globe.

Several technologies were used in order to enable the navigation activities, including observation of Sun, stars and planets and the development of navigation tools, such as the compass and astrolabe. These technologies evolved constantly, aiming to provide better accuracy.

In this scenario, the advent of Global Navigation Satellite Systems changed significantly the navigation techniques, starting with the development and launch in the 1970s of the Global Positioning System (NAVSTAR-GPS) (Novatel Inc., 2015). In the sequence of GPS launch, other initiatives such as GLONASS, Galileo and Beidou have emerged in the GNSS environment. These radio-navigation systems are composed of satellites orbiting the Earth and the basic operation principle is the measurement of the distance between the user and four satellites. In the following sections, the available GNSS systems and some of their

features will be briefly presented.

2.1.1 Global Positioning System (NAVSTAR-GPS)

The NAVSTAR-GPS developed by the US Department of Defense (DoD-USA), was conceived with the purpose to be the main navigation system of the North-American military forces. This system was originated after the fusion of two projects funded by the USA government for development of a global navigation system: *Timation* and *System 621 B*, under the responsibility of Navy and Air Force, respectively (MONICO, 2008). The original goals of the NAVSTAR-GPS were an instantaneous determination of position, velocity and time in a common reference system anywhere on or near the Earth (HOFMANN-WELLENHOF; LICHTENEGGER; WASLE, 2008).

Although this system was originally designed for military activities, it has been massively used by the civilian community in activities such as navigation, surveying and agriculture. (MONICO, 2008; HOFMANN-WELLENHOF; LICHTENEGGER; WASLE, 2008). The basic concept on GPS navigation is based on the range measurement between the user receiver and at least four satellites. By the knowledge of the satellite positions in a suitable reference system, it is possible to estimate the receiver position coordinates at the same reference system. From a geometric point of view, only three satellites would be necessary to allow the receiver position calculation. However, due to the non-synchronism between the receiver and satellites clocks, one more unknown is added to the position determination problem. It is common to have more than four satellites in view, which allows a better control of the quality of the solution. In this case, the least squares method is usually used to process the redundant measurements, leading to an improvement in the positioning solution (KAPLAN; HEGARTY, 2006; MONICO, 2008).

This system is divided into three segments: the spatial segment, control segment and user segment. The spatial segment is formed by the 24 satellites, orbiting in approximated altitude of 20.200 km and distributed in six equally spaced orbital planes with inclination of 55° with respect to the Equator. This configuration allows that at least four satellites are visible in any part of Earth at any time.

The system was declared operational on 27th April 1985. It has been improved and modernized since then. The driven force for modernization was both the military and civilian interests and requests. Also, the development of other systems such as the European Galileo and the Chinese BeiDou, stimulated the modernization process (HOFMANN-WELLENHOF; LICHTENEGGER; WASLE, 2008). The first launched satellites, the Block I, were prototypes and the last satellite of the Block was disabled at the end of 1995. The first and second generation of satellites ate the Blocks II and IIA, respectively. When the system was declared fully operational in 1995 all the satellites pertained to these blocks. A new generation of satellites is in production, the GPS III satellites, with the first launch expected to occur in 2018 (US Department of Defense, 2017).

The GPS now effectively operates as a 27-slot constellation with improved coverage in most parts of the world. As of October 17, 2017, there was a total of 31 operational satellites in the GPS constellation, not including the decommissioned, on-orbit spares (US Department of Defense, 2017).

The control segment consists of a master control station, monitor stations and ground antennas (HOFMANN-WELLENHOF; LICHTENEGGER; WASLE, 2008). The main purposes of the control segment is to continuously monitor and control the satellite system; determine the GPS time system; predict the satellite ephemeris; calculated the satellite clock corrections; and continuously update the navigation message of each satellite (MONICO, 2008; Novatel Inc., 2015). The ground antennas operations are under the master control station and are equipped to transmit data and commands to the satellites and to receive telemetry and ranging data from the satellites (HOFMANN-WELLENHOF; LICHTENEGGER; WASLE, 2008).

The user segment consists of the GPS receivers which use the information transmitted by the satellites to calculate user's three dimensional position and time (US Department of Defense, 2017). The equipments used to process the received GPS signals varies from smartphones and handheld receivers used by hikers, to sophisticated and specialized receivers used in surveying and mapping applications (Novatel Inc., 2015).

In the GPS, the satellites transmit messages using the same frequencies, but each satellite can be identified by its exclusive code. This technique is referred to as code division multiple access (CDMA). The GPS satellites can transmit information on the L1 and L2 carrier waves, both generated based on the fundamental frequency $f_o = 10.23$ MHz, multiplied by 154 and 120, respectively (MONICO, 2008). The satellites belonging to the Block IIF also transmit on the carrier frequency L5, which is obtained by multiplying the fundamental frequency by 115. The available GPS signals are presented in Table 2.1.

Table 2.1: Carrier frequency per signal in the GPS (MONICO, 2008).

Signal	Carrier frequency (MHz)
L1	1574.42
L2	1227.60
L5	1176.45

The L1 is modulated by the Coarse/Acquisition (C/A) code and the Precision (P) code, which are available for civilian and military/authorized users, respectively. The L2 signal is modulated by the P code and it was included in the system in order to allow users to correct automatically for the effects of both the range and the range rate errors due to the ionosphere (KLOBUCHAR, 1987). A new civilian signal referred to as L2C is available in the satellites belonging to Block IIR-M and later. This signal allows the direct measurement and correction of the ionospheric delay error, for a particular satellite, using civilian signals on both L1 and L2. The L5 signal was incorporated into the GPS in order to meet demanding requirements for safety-life transportation and other high-performance applications. The L5

is broadcast in a radio band exclusively reserved for aviation safety services (Novatel Inc., 2015; US Department of Defense, 2017).

2.1.2 GLONASS

Developed in the 1970s decade, by the former Union of Soviet Socialist Republics (URSS), the GLONASS (Global'naya Navigatsionnaya Sputnikovaya Sistema) was conceived with the purpose to provide 3-D positioning, velocity and time information under any climate condition in local, regional and global levels. Operated by the Russian military forces, the GLONASS is a military system, however, the Russian government has done several declarations offering the system for civilian uses (MONICO, 2008; HOFMANN-WELLENHOF; LICHTENEGGER; WASLE, 2008).

The first satellite was launched on October 12 (ALKAN; KAMMAN; SAHIN, 2005). In 1995 the system was declared operational, with a constellation of 24 satellites divided into three orbital planes and orbiting in approximated altitude of 19 100 km. Since it was declared operational, the number of available satellite decreased due to the lack of funding and the launching of new satellites, reaching only ten operational satellites in the end of 2006 (MONICO, 2008; HOFMANN-WELLENHOF; LICHTENEGGER; WASLE, 2008). In December 2017, the system had 24 operational satellites and one satellite in flight test phase (GLONASS information and Analysis Center, 2017)

Like GPS, each GLONASS satellite provides navigation signals in two L-band frequencies, the G_1 and G_2 signals. This notation enables a distinction from GPS carriers L_1 and L_2 . However, it is possible to find in the literature the notation L_1 and L_2 referring to GLONASS signals (HOFMANN-WELLENHOF; LICHTENEGGER; WASLE, 2008).

Although the GLONASS presents two carriers modulated by two binary codes, and the navigation messages, in this system each satellite transmit carrier signals at different frequencies, which corresponds to the frequency division multiple access (FDMA) (BORRE; STRANG, 2012). The G_1 frequencies are given by:

$$f_{G_1} = f_0 + k \times \Delta f_{G_1} \quad k = 0, 1, 2, \dots, 24 \quad (2.1)$$

where $f_0 = 1602$ MHz, $\Delta f_{G_1} = 0.5625$ MHz and k means the frequency number of the satellite. The carriers G_1 and G_2 have the following relation (SEEBER, 2003)

$$\frac{f_{G_1}}{f_{G_2}} = \frac{9}{7}. \quad (2.2)$$

In addition to the FDMA, the code division multiple access (CDMA) has been incorporated in the GLONASS since 2011. In the CDMA technology the satellites are allowed to transmit messages using the same frequencies. In this case, each satellite is identified by a different code. This technology is employed in all other GNSS (BORRE; STRANG, 2012).

2.1.3 Galileo

Developed by the European Space Agency (ESA), the European Commission (EC) and European industry, the Galileo is Europe own global navigation satellite system. The system is under civilian control and is interoperable with GPS and GLONASS (European Space Agency, 2017). In the full configuration, the system will have 24 satellites divided into three orbital planes and orbiting at an altitude of the order of 23 222 km. In addition to the 24 operational satellites, each orbital plane will have two spare satellites in case any operational satellite fail (European Space Agency, 2017).

The system was declared operational in 15 December 2016. Its first satellite was launched in December, 28, 2015 and by the end of December, 2017 the system presented fourteen operational satellites, three satellites under testing and four satellites under commissioning. The full constellation of 30 satellites is expected to be completed by 2020 (European Space Agency, 2017).

The Galileo signals are transmitted in three band frequencies (E5, E6 and E1) using four carriers (E5a, E5b, E6 and E1). Galileo carrier frequencies are shown in Table 2.2

Table 2.2: Carrier frequency per signal in the Galileo system (The European Commission, 2016).

Signal	Carrier frequency (MHz)
E1	1575.420
E6	1278.750
E5a	1176.450
E5b	1207.140

2.1.4 BeiDou

China, that in the past revolutionized the navigation with the development of the compass, has started the implementation of its GNSS BeiDou Navigation Satellite System (BDS). This implementation was divided into two phases: the initial phase, which provides regional coverage and the second phase that will provide global coverage. The first phase was declared operational on December 2012 and the second phase is expected to be concluded by the end of 2020 (Novatel Inc., 2015).

Different from the other systems presented previously, the space segment of the second phase of BDS will consist of five Geostationary Earth Orbit (GEO) satellites, three Inclined Geosynchronous Orbit (IGSO) satellites and twenty-seven Medium Earth Orbit (MEO) satellites, orbiting at altitudes of 35 787 km, 35 787 km and 21 528 km, respectively (Novatel Inc., 2015).

The system provides three types of service: public service for civilian use and free to users; licensed service available only to users who have a subscription and the restricted mil-

itary service (Novatel Inc., 2015). Some characteristics of the BeiDou signals are presented in Table 2.3.

Table 2.3: Signal characteristics in the BeiDou system (Novatel Inc., 2015).

Designation	Frequency (MHz)	Description
B1	1207.140	Provides both public and restricted service signals
B2	1278.750	Provides both public and restricted service signals
B3	1268.520	Provides restricted service signals only

2.2 The GNSS observables

The determination of the receiver position consists of the computation of its three coordinates with respect to the origin of a coordinate system. For this purpose, it is more convenient to use a coordinate system that rotates with the Earth, known as an Earth-centered Earth-fixed (ECEF) system (KAPLAN; HEGARTY, 2006). The basic GNSS observables that allow this computation are the pseudorange measured through the code; and the carrier phase (MONICO, 2008; TEUNISSEN; KLEUSBERG, 1996):

2.2.1 The pseudorange

Considering the ECEF coordinate system, the receiver positioning problem can be graphically represented as shown in Figure 2.1.

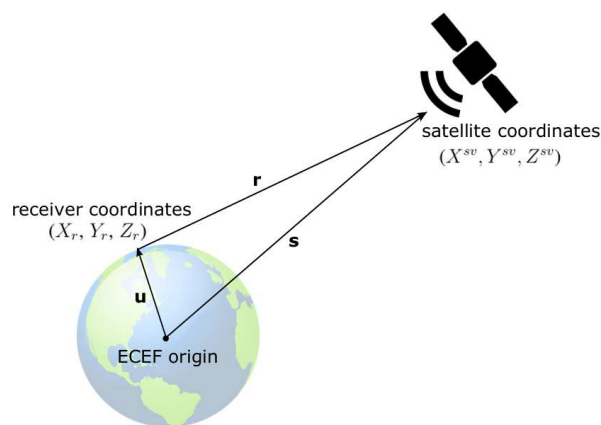


Figure 2.1: Receiver positioning calculation.

The purpose is to determine the vector \mathbf{u} , which is equivalent to determine the unknown position of the receiver with respect to the ECEF coordinate system origin. The receiver coordinates are represented by (X_r, Y_r, Z_r) . The vector \mathbf{r} represents the distance between the receiver and the satellite coordinates which are represented by (X^{sv}, Y^{sv}, Z^{sv}) . The vector \mathbf{s} represents the satellite location with respect to the origin of the coordinate system and can be

calculated using the ephemeris data transmitted to the user (KAPLAN; HEGARTY, 2006). The subscript r and the superscript sv refer to as the coordinates of the receiver and the satellite, respectively.

The measurement of the range between the satellite and the receiver antenna is based on the code generated in the satellite and a replica generated in the receiver. These codes are represented by $G^s(t)$ and $G_r(t)$, respectively. By measuring the delay between a particular transition in the code $G^s(t)$ and the replica code $G_r(t)$, the propagation time of the signal in the path from the satellite to the receiver is obtained. The receiver measures this delay by using the code cross-correlation. Figure 2.2 illustrates this principle (MONICO, 2008).

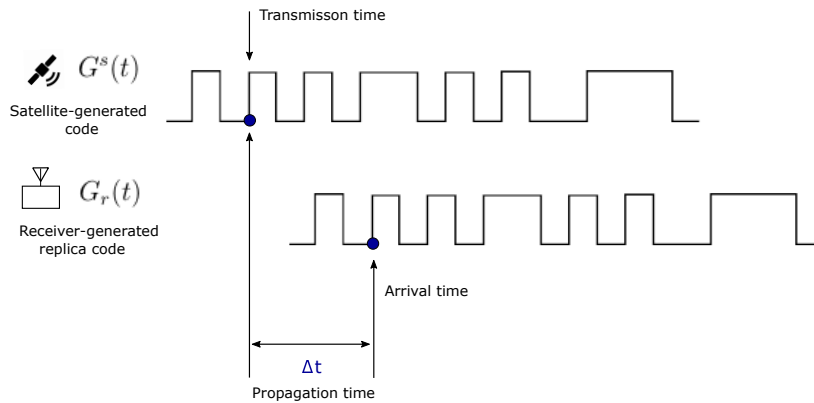


Figure 2.2: Use of replica code to determine the satellite code travelling time (MONICO, 2008; KAPLAN; HEGARTY, 2006) (*adapted*).

The replica generated in the receiver is shifted until a high correlation between the signal transmitted by the satellite and the replica generated at the receiver is reached. If the satellite clock and the receiver clocks were perfectly synchronized, the correlation process would yield true propagation time. By multiplying the propagation time Δt by the speed of light, c , the true (i.e. geometric) satellite-to-user range would be obtained. However, this is an ideal scenario that considers the clocks synchronism. In general, the satellite and receiver clocks are not synchronized (KAPLAN; HEGARTY, 2006).

The receiver clock generally presents a *bias* with respect to the system time. Furthermore, although the time and frequency generation of the satellite is based in high accuracy atomic clocks (cesium and rubidium), the satellite clock also presents an offset with respect to the system time as well (KAPLAN; HEGARTY, 2006). Thus, due to the clock errors, the range calculated by using the correlation process is denoted as the pseudorange P (KAPLAN; HEGARTY, 2006).

The GNSS satellites have high precision atomic clocks operating in the satellite time system (t^{sv}), to which all generated and transmitted signals are referenced. The receivers, in general, have lower quality oscillators that operate in the receiver time system (t_r), to which the received signals are referenced. For the GPS, these time systems, t^{sv} and t_r , can

be related to the GPS time system (t_{GPS}) according to Equation (2.3) (MONICO, 2008).

$$\begin{aligned} t_{GPS^{sv}} &= t^{sv} - dt^{sv} \\ t_{GPS_r} &= t_r - dt_r, \end{aligned} \quad (2.3)$$

where, dt^{sv} is the satellite clock error with respect to the GPS time at the instant t^{sv} and dt_r is the receiver clock error with respect to the GPS time at the instant t_r . The subscripts and superscripts refer the parameters related to the receiver and satellite, respectively (MONICO, 2008).

The pseudorange (P_r^{sv}) is obtained by multiplying the velocity of light by the difference between the time t_r registered at the receiver in the instant of signal reception and the time t^{sv} , registered at the satellite in the instant of signal transmission. Using the correlation process to obtain the propagation time, one can obtain the following expression for the pseudorange (MONICO, 2008; TEUNISSEN; KLEUSBERG, 1996):

$$P_r^{sv} = c(t_r - t^{sv}) = c(t_{GPS_r} - t_{GPS^{sv}}) + c(dt_r - dt^{sv}) + \varepsilon_P = c\tau_r^{sv} + c(dt_r - dt^{sv}) + \varepsilon_P, \quad (2.4)$$

where τ_r^{sv} is the propagation time of the signal, counted from its generation at the satellite until the correlation at the receiver, c is the velocity of light at vacuum, and ε_P is the pseudorange measurement error (MONICO, 2008).

The propagation time τ_r^{sv} multiplied by the velocity of light on vacuum does not result in the geometric distance ρ_r^{sv} between the antenna of the satellite and the receiver due to other sources of errors, such as the propagation effects of the atmosphere (e.g tropospheric and ionospheric delays) and multipath. Thus, a more complete form for Equation (2.4) is (MONICO, 2008):

$$P_r^{sv} = \rho_r^{sv} + c(dt_r - dt^{sv}) + T_r^{sv} + I_r^{sv} + dm_r^{sv} + \varepsilon_P, \quad (2.5)$$

where

ρ_r^{sv} is the geometric distance from the satellite to the receiver;

T_r^{sv} is the tropospheric delay (in meters);

I_r^{sv} is the ionospheric delay (in meters);

dm_r^{sv} is the multipath effect.

The coordinates of the receiver and the satellite are implicit in the term ρ_r^{sv} as presented in Equation (2.6).

$$\rho_r^{sv} = \sqrt{(X^{sv} - X_r)^2 + (Y^{sv} - Y_r)^2 + (Z^{sv} - Z_r)^2}. \quad (2.6)$$

Applying (2.6) in (2.5), the pseudorange expression with the receiver and satellites coor-

coordinates presented explicitly is given by:

$$P_r^{sv} = \sqrt{(X^{sv} - X_r)^2 + (Y^{sv} - Y_r)^2 + (Z^{sv} - Z_r)^2} + c(dt_r - dt^{sv}) + T_r^{sv} + I_r^{sv} + dm_r^{sv} + \varepsilon_P. \quad (2.7)$$

Given the spatial satellite distribution on GNSS orbits, it is very common to have pseudorange measurements from more than four satellites. Thus, an overdetermined system is available to calculate the variables of interest (receiver coordinates and receiver clock bias). This observable is nonlinear with respect to the satellites and receivers coordinates. More information about the linearization of Equation (2.7) can be found in the Appendix B.

2.2.2 The carrier phase

The carrier phase observable (Φ_r^{sv}) is more precise than the pseudorange and it is the most common observable employed in applications that require a high accuracy.

Ideally, the carrier phase observable would be equal to the sum of the total number of full carrier cycles and the fractional cycles between the antennas of a satellite and a receiver at any instant (LANGLEY, 1996). However, only the fractional carrier phase can be measured by the receiver. The number of full cycles N_r^{sv} is unknown and have to be estimated with the other unknowns (TEUNISSEN; KLEUSBERG, 1996; LANGLEY, 1996). This observable is presented in Equation (2.8)

$$\Phi_r^{sv} = \rho_r^{sv} + c(dt_r - dt^{sv}) + \lambda N_r^{sv} - I_r^{sv} + T_r^{sv} + dm_r^{sv} + \varepsilon_{\Phi_r^{sv}}, \quad (2.8)$$

where λ is the wavelength of the carrier, N_r^{sv} is the ambiguity and $\varepsilon_{\Phi_r^{sv}}$ is the carrier phase error. It is important to note that in the carrier phase observable the ionospheric effect is negative, whereas in the pseudorange this effect is additive (MONICO, 2008). More information about this fact is presented in the section 2.3.4.

2.3 GNSS observables error sources

The GNSS observable are subject to random, systematic and gross errors, like all observables involved in measurements activities. Systematic errors can be parametrized or reduced (or even eliminated) by suitable techniques. Random errors, however, do not present functional relation with the measurements and generally are the remaining errors in the observations, after all systematic and gross errors are reduced or minimized (MONICO, 2008). In the next sections, some of the common errors in GNSS observations will be presented, specifically the clock errors, multipath, tropospheric and ionospheric effects. Other sources of errors can be found in Appendix E.

2.3.1 Clock errors

2.3.1.1 Satellite clock errors

The satellite atomic clocks are monitored by the GNSS control segment and are very precise, however, they do not work perfectly synchronized with the GNSS reference time. From the data provided in the navigation message, it is possible to obtain the satellite clock correction dt^{sv} for a satellite sv by the following second order polynomial (MONICO, 2008):

$$dt^{sv}(t) = a_0 + a_1(t^{sv} - t_{oc}) + a_2(t^{sv} - t_{oc})^2 + \Delta t_R, \quad (2.9)$$

where

δ^{sv} is the satellite clock error on the instant t of the GNSS time scale;

t^{sv} is the satellite reference epoch;

t_{oc} is the reference time of clock data;

a_0 is the satellite *clock offset* coefficient, in seconds;

a_1 is the satellite *clock drift* coefficient;

a_2 is the satellite *clock frequency drift* coefficient;

$\Delta t_R = -2 * X * \dot{X} / c^2$ is the correction for the relativistic effect on the satellite *clock* (X , \dot{X} and c are the position of the satellite, its velocity and the velocity of light, respectively).

2.3.1.2 Receiver clock errors

Differently from satellite clock errors, the receiver clock corrections in point positioning applications need to be performed by the user. A brief description about point positioning is presented in Appendix F. This procedure is done by the estimation of an additional parameter that refers to the receiver clock synchronization error for every observation epoch. Using this procedure allows for the employment of small and inexpensive oscillators in the receivers, such as quartz crystal oscillators (WEINBACH; SCHÖN, 2011; MONICO, 2008).

However, this type of clock estimation presents some associated drawbacks, such as the need to observe at least four-satellites at the same time to determine the receiver position (three coordinates and one receiver clock off-set). Another point that has to be taken into account is the degradation of the vertical position accuracy. This effect occurs due to the asymmetry on the observations, since only satellites in the hemisphere above the horizon are observed. This fact leads to a mathematical correlation between receiver clocks, troposphere parameters and station height estimates (WEINBACH; SCHÖN, 2011; M.ROTHACHER; G.BEUTLER, 1998).

In the relative positioning, the clock errors are almost eliminated, being not necessary to use highly stable clocks for the majority of applications. However, the simultaneity of the

observations has to be carefully taken into account. For high accuracy results, the receiver clock error in the relative positioning for each receiver involved in this technique has to be known to 1 microsecond ($1\mu s$) with respect to the time system and the differences between them cannot exceed 1 ms (MONICO, 2008).

2.3.2 Multipath

Considered one of the most significant sources of errors in satellite-based navigation systems, the multipath may introduce errors in positioning calculations that could jeopardize high-precision applications (CLOSAS; FERNÁNDEZ-PRADES; FERNÁNDEZ-RUBIO, 2009). Multipath is the phenomena in which the signal reaches the receiver by multiple paths due to reflection and diffraction (BRAASCH, 1996). In other words, the receivers will get the signal which reaches directly the antenna and also signals reflected in the surfaces nearby. An illustration of the multipath is presented in Figure 2.3.

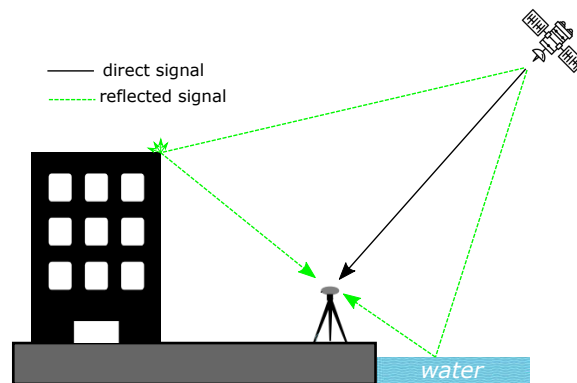


Figure 2.3: Multipath illustration.

The signal received can present distortions on the carrier-phase and on modulation of the carrier, and since the geometrical features in each place changes in an arbitrary way, there is no model available to mitigate multipath. However, some techniques can be applied in order to reduce this effect, including the use of antennas designed to suppress low-elevation-angle signals, such as the *choke ring* and *pinwheel*, and its well-placement in suitable sites (ARBESSER-RASTBURG; ROGERS, 2013; BISHOP; KLOBUCHAR; DOHERTY, 1985; MONICO, 2008).

2.3.3 Tropospheric effect

The troposphere is the layer of atmosphere that starts from the Earth surface and extends until 50 km, approximately. It behaves like a non-dispersive medium for frequencies below 30 GHz, i.e. the refraction of the transmitted signal does not depend on the signal frequency. It depends only on the thermodynamic properties of the air (MONICO, 2008). The specific tropospheric effects on the GNSS L-band signals include tropospheric attenuation, tropo-

spheric scintillation and tropospheric delay (SPILKER, 1996). This section presents a quick overview of the tropospheric effect on GNSS signals. Additional information can be found in the Appendix C.

2.3.3.1 Tropospheric attenuation

The tropospheric attenuation varies for each frequency and corresponds to the reduction of power of the electromagnetic wave due to the elements that compose the atmosphere (MONICO, 2008; SPILKER, 1996). In the 1-2 GHz frequency band, it is dominated by oxygen attenuation. For a satellite at zenith, the attenuation is about 0.035 dB. The effects of water vapor, rain and nitrogen attenuation is negligible at GNSS frequency bands (SPILKER, 1996). For GNSS signals it is not recommended to use observations obtained at elevation angles lower than 5° , since the ray path to the satellite penetrates the lower troposphere in a more nearly horizontal direction leading to a higher signal attenuation (SPILKER, 1996). In practice, it is common to use elevation angles higher than 15° , commonly referred to as *elevation mask* (MONICO, 2008). More information about the tropospheric attenuation as a function of the elevation angle can be found in Appendix C.

2.3.3.2 Tropospheric delay

The delay on the GNSS signals when traveling through the troposphere is caused mainly by the neutral hydrostatic atmosphere (composed of dry gases), corresponding to 90% of the total effect. The remaining 10% depends on the water vapor 3D distribution (non-hydrostatic component), which is hard to estimate, due to its high temporal and spatial variations (HADAS et al., 2013; SAPUCCI, 2001).

The hydrostatic component of the tropospheric delay corresponds to 2.3 m on zenith and varies with temperature and local atmospheric pressure. Since its variation is small (in the order of 1% over several hours), this component is predicted with reasonable precision. The wet effect, caused by the atmospheric water vapor influence, is less than the hydrostatic component, varying from 1 to 35 cm in zenith, corresponding to 10% of the total tropospheric delay (SEEBER, 2003; MONICO, 2008). However, although its low effect, its variation is considerable, reaching 20% in a few hours, making impossible its prediction with good precision, even when there is availability of superficial humidity measurements (MONICO, 2008).

In general, the models to estimate tropospheric delay in the path between the receiver antenna r and the satellite sv are presented as

$$T_r^s = 10^{-6} \int N_T ds. \quad (2.10)$$

The troposphere refractivity is given by $N_T = (n - 1) \times 10^6$, where n is the refractive

index. The hydrostatic and wet components of the tropospheric delay can be written as the product between the zenith delay and a mapping function which relates the vertical delay with the slant delay. Therefore, the tropospheric delay can be expressed as

$$T_r^s = T_{ZH} \times mh(\epsilon) + T_{ZW} \times mw(\epsilon), \quad (2.11)$$

where T_{ZH} and T_{ZW} corresponds, respectively, to the hydrostatic and wet components of the tropospheric delay in the zenith direction. The $mh(\epsilon)$ and $mw(\epsilon)$ values are the corresponding mapping functions used to convert the zenith delay to slant delay (MONICO, 2008).

In order to determine the tropospheric delay T_r^s it is necessary to obtain the refractivity N_T . By the knowledge of N_T and its hydrostatic (N_H) and wet (N_W) components it is possible to determine the terms T_{ZH} and T_{ZW} , and consequently, T_r^s by using the mapping functions. The determination of the refractivity along the signal path is almost impossible (SAPUCCI, 2001). This is the reason why several models have been developed to describe the behavior of this variable. More information about the tropospheric models and the mapping functions can be found in Appendix C.

2.3.4 Ionospheric effect

The ionosphere is the Earth atmosphere region where ionizing radiation causes the existence of electrons in an amount that affects radio waves propagation (LANGLEY, 1992). The energy radiated from the sun at ultraviolet and X-ray wavelengths is the primary force of ionosphere formation. This force ionizes gaseous atoms and molecules in the atmosphere, producing positively charged ions and negatively charged free electrons (WEBSTER, 1993). In order to describe the amount of charged particles in the ionosphere, the term *electron density* (N_e) is used (HABARULEMA et al., 2009).

The ionosphere is composed of four regions, D , E , F_1 and F_2 , named in order of increasing height. Each region presents particular electron density features (KLOBUCHAR, 1996). Table 2.4 presents some characteristics of the ionospheric regions, such as the height of each layer, and information about its impact on GNSS signals.

As described in Section 2.3.3.2, the tropospheric range error at zenith is generally between two to three meters. The ionospheric range error, on the other hand, can vary from only a few meters to many tens of meters at the zenith (KLOBUCHAR, 1996). When compared to the tropospheric effect, the variability of the ionospheric effect is much larger and it is more difficult to model. Furthermore, ionosphere can have significant effects on GNSS, such as: group delay of the signal modulation, or absolute range error; carrier phase advance, relative range error; Doppler shift, or range-rate errors; Faraday rotation of linearly polarized signals; refraction or bending of the radio wave; distortion of pulse waveforms; signal amplitude fading or amplitude scintillation; and phase scintillations. (KLOBUCHAR, 1996;

Table 2.4: Characteristics of the different regions of the ionosphere (KLOBUCHAR, 1996).

Region	Approx. Height (km)	Remark
D	50 to 90	Causes absorption of radio signals at frequencies up to low VHF band. It has no measurable effect on GNSS
E	90 to 140	Normal E region, caused by solar soft x-rays. It has minimal effect on GNSS. Intense E region, caused by solar particle precipitation in aurora region, might cause minimal scintillation. Sporadic E (still of unknown origin). It is very thin. The effect on GNSS frequencies is neglected.
F_1	140 to 210	It has a highly predictable density from known solar emissions. Its electron density nicely merges into the bottomside of the F2 region.
F_2	210 to 1000	It is the most dense and it has the highest variability. The peak of electron density generally varies from 250 to 400 km, but it can differ at extreme conditions. This region, with to some extension the F_1 , cause most of the problems for radio-wave propagation at GNSS frequencies.

HOQUE; JAKOWSKI, 2012).

The interaction between the GNSS radio signals and the ionospheric plasma is one of the major reasons for the limited accuracy and vulnerability in GNSS positioning solutions or time estimation (HOQUE; JAKOWSKI, 2012). Even at this relatively high frequency, the Earth ionosphere can retard radio waves from their velocity in free space by more than 300ns on a worst case basis, corresponding to range errors of 100m (JAKOWSKI et al., 2011).

An electromagnetic wave is not affected when traveling through a static electric or magnetic field in a linear medium such as vacuum. However, traveling through the ionosphere, which is a dispersive medium (i.e. the velocity of the electromagnetic wave is a function of its frequency), different aspects cause variation on the polarization, propagation speed and signal power (BORRE; STRANG, 2012; HOQUE; JAKOWSKI, 2015; HOQUE; JAKOWSKI, 2012). In the context of L band signals of GNSS, the ionospheric refraction, proportional to the TEC value, introduces most of the delay and may cause link-related range errors in the positioning system of up to 100 m (HOQUE; JAKOWSKI, 2015). Three different dynamic processes contribute to strong temporal and spatial variation of TEC in the equatorial and low-latitude sectors: the equatorial ionization anomaly, post-sunset plasma enhancement and evening plasma bubbles, leading to an even worst scenario in these regions (TAKAHASHI et al., 2014).

The propagation of an electromagnetic wave passing through the ionosphere is quantitatively described by the refractive index of the ionosphere (HOQUE; JAKOWSKI, 2012). For radio waves with frequency f greater than 100 MHz, the phase refractive index n and the group refractive index n_{gr} derived from the Appleton-Hartree equation are given by Equa-

tions (2.12) and (2.13).

$$n = 1 - \frac{f_p^2}{2f^2} \pm \frac{f_p^2 f_g \cos\Theta}{2f^3} - \frac{f_p^2}{4f^4} \left[\frac{f_p^2}{2} + f_g^2(1 + \cos^2\Theta) \right], \quad (2.12)$$

$$n_{gr} = 1 + \frac{f_p^2}{2f^2} \mp \frac{f_p^2 f_g \cos\Theta}{f^3} + \frac{3f_p^2}{4f^4} \left[\frac{f_p^2}{2} + f_g^2(1 + \cos^2\Theta) \right], \quad (2.13)$$

in which Θ is the angle between the wave propagation direction and the geomagnetic field vector \mathbf{B} and f_p and f_g are the plasma frequency and gyro frequency given by

$$\begin{aligned} f_p^2 &= N_e e^2 / (4\pi^2 \varepsilon_0 m), \\ f_g &= eB / (2\pi m), \end{aligned} \quad (2.14)$$

where ε_0 is the free space permittivity, B is the geomagnetic induction, and e , N_e , m are the electron charge, density and mass, respectively (HOQUE; JAKOWSKI, 2012). The expression presented in Equation (2.13) can be obtained by the relationship $n_{gr} = n + f(dn/df)$ (HOQUE; JAKOWSKI, 2012; HARTMANN; LEITINGER, 1984; APPLETON, 1932). The positive and negative signals in the equations \pm and \mp are related with the polarization of the wave, which means that the (+) sign represents the refractive index for left-hand circularly polarized wave, whereas the (-) represents the right-hand circularly polarized wave. (HOQUE; JAKOWSKI, 2012; HARTMANN; LEITINGER, 1984). The GPS signals are transmitted in right-hand circular polarization (DoD, 2012).

By analyzing Equations (2.12) and (2.13) one can note that the phase refractive index is less than the unity resulting in a phase velocity that is greater than the speed of light in the vacuum (i.e., phase advance). Thus, when GNSS signals propagate through the ionosphere, the carrier-phase experiences an advance and the code experiences a group delay. The carrier-phase pseudoranges are measured too short and the code pseudoranges are measured too long compared to the geometric range between a satellite and a receiver (HOQUE; JAKOWSKI, 2012).

2.3.4.1 Total Electron Content

By using Equations (2.12) and (2.13) and assuming a right-hand circularly polarized signal, the ionospheric group delay d_{Igr} and the carrier phase advance d_I , written in units of length, in the path s from the satellite to the receiver are given by (HOQUE; JAKOWSKI, 2012)

$$d_{Igr} = d_{Igr}^{(1)} + d_{Igr}^{(2)} + d_{Igr}^{(3)} = \int_s (n_{gr} - 1) ds = \frac{p}{f^2} + \frac{q}{f^3} + \frac{u}{f^4}, \quad (2.15)$$

$$d_I = d_I^{(1)} + d_I^{(2)} + d_I^{(3)} = \int_s (1 - n) ds = \frac{p}{f^2} + \frac{q}{2f^3} + \frac{u}{3f^4}, \quad (2.16)$$

in which the terms $d_{Igr}^{(1)} / d_I^{(1)}$, $d_{Igr}^{(2)} / d_2^{(2)}$, and $d_{Igr}^{(3)} / d_I^{(3)}$ are the first, second and third order ionospheric group delays / phase advances, respectively. The coefficients p , q and u are given by

$$p = 40.3 \int_s N_e ds, \quad (2.17)$$

$$q = 2.2566 \times 10^{12} \int_s N_e B \cos \Theta ds, \quad (2.18)$$

$$u = 2437 \int N_e^2 ds + 4.74 \times 10^{22} \int_s N_e B (1 + \cos^2 \Theta) ds. \quad (2.19)$$

The result of the integration of N_e along the signal path $\int_s N_e ds$ is known as the slant Total Electron Content (slant TEC). This parameter is defined as the amount of electrons in a column of cross sectional area of 1 m^2 along the path of the signal through the ionosphere and is expressed in TECU (1 TECU is equivalent to $1 \times 10^{16} \text{ electrons/m}^2$) (HABARULEMA et al., 2009). Considering the GPS, 1 TECU corresponds to a ionospheric delay of 0.163 m on L1 and 0.267 m on L2 signals (DYRUD et al., 2008).

When the GNSS signals are transmitted in two different signals, the non-dispersive effects, such as, tropospheric delay, satellite and receiver clocks biases affects equally both frequencies. The ionospheric effect, however, affects differently each frequency. Therefore, by differencing the code/carrier-phase pseudoranges measurements of the two frequencies it is possible to estimate the TEC along the path between the satellite and the receiver, as shown in Equations 2.20 and 2.21 (HOQUE; JAKOWSKI, 2012).

$$TEC = \frac{f_1^2 f_2^2}{40.3(f_1^2 - f_2^2)} [(P2 - P1) + noise_{P1-P2}], \quad (2.20)$$

$$TEC = \frac{f_1^2 f_2^2}{40.3(f_1^2 - f_2^2)} [(\Phi1 - \Phi2) + B_{ambiguity} + noise_{\Phi1-\Phi2}], \quad (2.21)$$

where $P1$ and $P2$ are pseudoranges obtained using the L1 and L2 signals, respectively; $\Phi1$ and $\Phi2$ are the carrier-phase measurements, the $noise_{P1-P2}$ and $noise_{\Phi1-\Phi2}$ are noises (e.g thermal noises, etc) in the code and carrier-phase combinations, $B_{ambiguity}$ is the carrier-phase constant ambiguity which is equal to $\lambda_2 N2 - \lambda_1 N1$, with λ_1 and λ_2 corresponding to the wavelengths and $N1$ and $N2$ corresponding to the integer ambiguities on frequencies f_1 and f_2 . The terms corresponding to other effects, such as, multipath and inter-frequency satellite and receiver biases are not presented in Equations 2.20 and 2.21 for simplicity reasons (HOQUE; JAKOWSKI, 2012).

The ionospheric effect is still one of the major error sources of GNSS positioning for single-frequency users (LIU et al., 2016). The first order term presented in Equations (2.15) and (2.16) includes about 99% of the total ionospheric effect. Therefore, if the frequency and the related slant TEC are known, the first order propagation effect can easily be computed and corrected. To compensate this effect, several approaches can be adopted. The

ionosphere free combination (IF) takes advantages of the dispersive character of the ionosphere to eliminate this first-order effect (ELMAS et al., 2011). This approach is based on multi-frequency satellite-receiver communication. Single-frequency receivers, however, can not take advantage of IF. In this case, observations need to be corrected by the use of external models in order to eliminate the first order ionospheric error contribution (PAPARINI et al., 2016). Information about some models available are presented in the next section.

2.3.4.2 Ionospheric models

The ionospheric models available are classified into empirical and physics based. The empirical models are developed based on long-term measurements and statistical analysis, whereas the physics-based models rely on a complete understanding of the underlying physics and commonly starts from an empirical model of the neutral atmosphere (MITCHELL, 2013). Some common ionospheric models are presented below:

- a) *Klobuchar* - The Klobuchar model received the name of his proposer John A. Klobuchar and was developed at the Air Force Geophysics Laboratory, U.S. This is a very known simple mathematical model that represents the ionosphere using a heuristic approach to finding a fit to some known observations. The purpose in designing this model is to provide a simple and fast ionospheric time-delay correction algorithm for the single-frequency users of GPS. The idea is to use a few coefficients, leading to a lesser computational effort to compensate for 50% of the ionospheric time-delay (KLOBUCHAR, 1987; MITCHELL, 2013).

The algorithm, incorporated in the GPS system for single-frequency users, consists of a cosine representation of the diurnal curve, allowed to vary in amplitude and in period, with user latitude. Its shape is determined by using eight coefficients, transmitted as part of the satellite message. Figure 2.4 presents an example in which a positive half cosine-shaped curve has been made to fit a typical monthly average TEC diurnal variation from an station in Jamaica using the referred model (KLOBUCHAR, 1987).

- b) *NeQuick* - The NeQuick model is a quick-run model for transionospheric applications developed at the Abdus Salam ICTP Aeronomy and Radiopropagation Laboratory, Italy, with collaboration of the University of Graz Institute for Geophysics, Astrophysics and Meteorology, Austria. This model provides electron concentration at the given location in space and time (N_e) and its basic inputs are position, time, and solar flux (or sunspot number). The NeQuick package allows to compute the electron density along any-ray path and by numerical integration provides the TEC to the user. Its original version has been used by the European Navigation Overlay Service (EGNOS) of the European Space Agency (ESA) for system assessment analysis.

One very important use of the NeQuick model is its adoption as the model for ionospheric corrections for single frequency users on the GALILEO system. To be used in

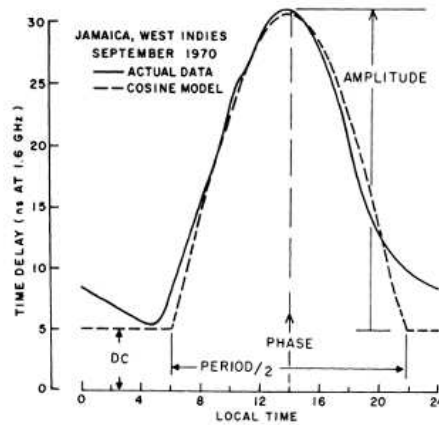


Figure 2.4: Example of actual monthly average time-delay data, along with cosine model fit to data (KLOBUCHAR, 1987).

GALILEO the model is driven by an *effective ionization level* (A_z), defined as:

$$A_z = an_0 + an_1\mu + an_2\mu^2, \quad (2.22)$$

where μ is the modified dip coordinate (*modip*). The coefficients an_0, an_1, an_2 allow the A_z computation and are broadcast to the user.

The improvements made in the original NeQuick model resulted in a new version of it, named NeQuick 2. Its on-line version is available at:

<<https://t-ict4d.ictp.it/nequick2/nequick-2-web-model>>.

- c) *Global Ionospheric Map (GIM)* - The Global Ionospheric Maps are provided in the IONEX format (*IONosphere map Exchange format*) with the vertical TEC (vTEC) values presented in grids with spatial resolution of $5^\circ \times 2.5^\circ$ longitude and latitude, respectively, distributed from 87.7° (North) to 87.5° (South) in latitude and 180° (West) to 180° (East) in longitude. These GIMs are provided by the centers that contribute to the IGS, such as, CODE (Center for Orbit Determination), ESA (European Space Agency), JPL (Jet Propulsion Laboratory), and UPC (Polytechnic University of Catalonia).

The Global Ionospheric maps also provide the *Differential Code Biases (DCB)* for the codes P1-P2 and P1-C1. The precision of the vTEC provided by the map varies with the location, amount of GNSS stations used, level of ionospheric activity, and others. In addition to the GIMs, the Regional Ionospheric Maps (RIMs) have been investigated and developed by the scientific community. In Brazil, one can cite the developments done under the EMBRACE project (Brazilian Study and Monitoring of the Space Weather) of the INPE (National Institute for Spatial Research), and the RIM Mod_Ion developed at the Cartography Department of UNESP (São Paulo State University). The model allows estimating the error due to ionosphere in the carrier L_1 and also TEC. By using the coefficients provided by this model it is possible to correct the L_1 GPS observables (ROCHA; MARQUES; GALERA, 2015; AGUIAR, 2005).

In addition to the models presented herein, some modeling activities based on Artificial Neural Networks (NNs) have been carried out. The use of NNs has provided good results in the modeling of the regional TEC, being capable of recovering TEC values with good performance (HABARULEMA et al., 2009; LEANDRO; SANTOS, 2007). This fact is related to the abilities of a NN to learn, generalize and adapt to different patterns of input/output sets with nonlinear behavior (HAYKIN, 1999).

In this framework, this manuscript presents a NN regional model to estimate TEC behavior based on GNSS measurements in three different regions in Brazil (Northeast, Central-West and South regions). The next section presents general information and concepts about the Neural Networks, emphasizing the Multilayer Perceptron (MLP) structures since this class of NN formed the basis for the vTEC estimations performed in this work.

2.4 Artificial Neural Networks

An Artificial Neural Network (ANN) is a massively parallel distributed processor made up of simple processing units that has a natural propensity for storing experiential knowledge and making it available for use (HAYKIN, 2009). This network is inspired in biological neuronal systems (HAYKIN, 2009) and has been employed in studies in different fields, including applications on ionospheric modelling (HABARULEMA et al., 2009; LEANDRO; SANTOS, 2007; WILLISCROFT; POOLE, 1996).

The ANNs are formed by the massive interconnection of information-processing units called *artificial neurons*. Figure 2.5 presents the block diagram which depicts the model of an *artificial neuron*.

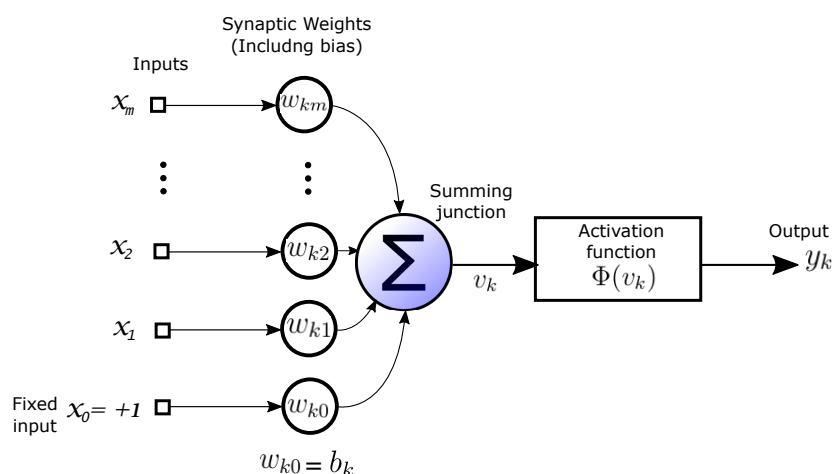


Figure 2.5: Artificial neuron model (HAYKIN, 2009) (*adapted*).

The three basic elements of the neuron model are presented in Figure 2.5: a set of *synapses* (or connecting links), an *adder* and an *activation function*. Each *synapse* is characterized by a *weight* or *strength* of its own. Specifically, a signal x_j at the input of synapse j connected to neuron k is multiplied by the synaptic weight w_{kj} . The *adder* constitutes a

linear combiner in which the input signals weighted by the respective synaptic strengths are summed. The role of the *activation function* is also to limit the amplitude of a neuron. It limits the permissible amplitude range of the output signal to some finite value. The model presented in Figure 2.5 also includes an externally applied *bias*, b_k , which has the effect of increasing or lowering the net input of the activation function, depending on whether it is positive or negative, respectively (HAYKIN, 2009).

2.4.1 Rosenblatt perceptron

The Rosenblatt perceptron is built around a single neuron and is the simplest form of a neural network. It consists of a neuron as the one presented in Figure 2.5 with adjustable free parameters (synaptic weight and bias). For a given set of inputs, the neuron produces an output equal to +1 if the activation function input is positive, and -1 if it is negative. The algorithm used to adjust the free parameters of this type of NN appeared for the first time in a learning procedure developed by Rosenblatt (1958, 1962) for his perceptron brain model (HAYKIN, 2009).

The goal of the single-neuron perceptron is to classify the input set into one of two classes \mathcal{C}_1 or \mathcal{C}_2 . Depending on the output of the network, +1 or -1, the point represented by inputs x_1, x_2, \dots, x_m will be assigned to class \mathcal{C}_1 or \mathcal{C}_2 , respectively.

To develop insight into the behavior of a pattern classifier, it is customary to plot a map of the decision regions in the m -dimensional signal space spanned by the m input variables x_1, x_2, \dots, x_m . In the single-neuron perceptron, there are two decision regions separated by a hyperplane, which is defined by

$$\sum_{i=1}^m w_i x_i + b = 0. \quad (2.23)$$

For the case of two input variables x_1 and x_2 , the hyperplane is given by $w_1 x_1 + w_2 x_2 + b = 0$. In this case, the decision boundary (*hyperplane*) takes the form of a straight line. A point (x_1, x_2) which lies below the line can be classified as \mathcal{C}_2 and a point (x_1, x_2) above the line can be classified as \mathcal{C}_1 . Figure 2.6 illustrates this decision boundary. It is possible to see that the *bias* b plays the role of shifting the decision boundary away from the origin (HAYKIN, 2009).

2.4.2 Multilayer Perceptron

The Rosenblatt perceptron, which is basically a single-layer neural network, is limited to the classification of linearly separable patterns (HAYKIN, 2009). To overcome this limitation a neural network structure that can be used is the Multilayer Perceptron (MLP). This structure contains one or more layers that are hidden from both the input and output and

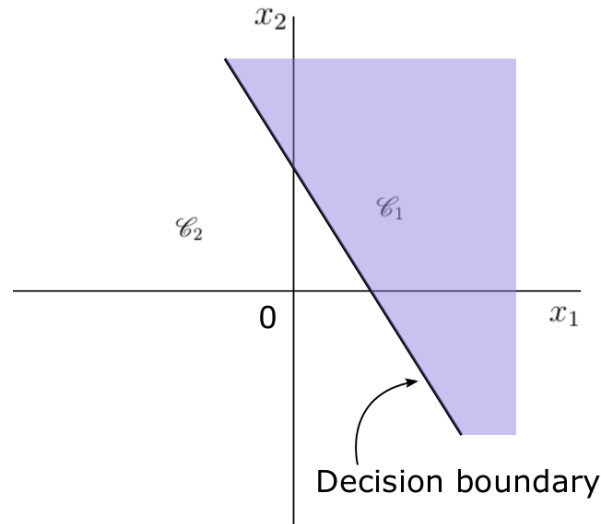


Figure 2.6: Hyperplane as a decision boundary for a two-dimensional, two-class pattern-classification problem (HAYKIN, 2009) (*adapted*).

presents a high degree of connectivity between the neurons. In addition, the neuron model used in the network includes a nonlinear activation function that is differentiable (HAYKIN, 2009). The sigmoid functions are the most common type of activation functions applied in neural networks. Its name is derived from the shape of its graph which is presented in a "S-shape" (HAYKIN, 2009). One example of sigmoid function is the logistic sigmoid function described by

$$\Phi(v) = \frac{1}{1 + \exp(-av)}, \quad (2.24)$$

where a is the slope parameter of the function.

In some cases, it is desirable that the activation function assumes positive and negative values, from -1 to $+1$. This kind of activation function may yield practical benefits over the logistic function described in Equation (2.24). In this case, the hyperbolic tangent function described in Equation (2.25) can be used.

$$\Phi(v) = \frac{\exp(v) - \exp(-v)}{\exp(v) + \exp(-v)}. \quad (2.25)$$

2.4.2.1 Function approximation

The MLP structure can be employed in the function approximation task (HAYKIN, 2009). Consider a nonlinear input-output mapping described by

$$\mathbf{d} = \mathbf{f}(\mathbf{x}) \quad (2.26)$$

where \mathbf{x} and \mathbf{d} are the input and output vectors, respectively. The vector-valued function \mathbf{f} is assumed to be unknown. In order to compensate the lack of knowledge about \mathbf{f} , the set of

training samples \mathcal{T} composed of labeled examples is given

$$\mathcal{T} = \{(\mathbf{x}_i, \mathbf{d}_i)_{i=1}^N\}. \quad (2.27)$$

The requirement is to develop a neural network that approximates the unknown function \mathbf{f} in a certain way that the following Euclidean distance over all inputs, such that

$$\|\mathbf{F}(\mathbf{x}) - \mathbf{f}(\mathbf{x})\| < \epsilon, \quad \text{for all } \mathbf{x}, \quad (2.28)$$

where ϵ is a small positive number and \mathbf{F} describes the input-output mapping actually realized by the NN. Considering that the size N of the training set \mathcal{T} is large enough and the network has the suitable number of free parameters, then the approximation error can be made small enough for the proposed task. This sort of approximation can be seen as a candidate for the supervised learning procedure, which requires the availability of a target response for performing a specific input-output mapping by minimizing a cost function of interest, with \mathbf{x}_i being the input vector, \mathbf{d}_i being the desired response, in which the index i represents the i -th input-output pair of the training set. In this function approximation scenario, a MLP containing only one hidden layer is capable to approximate any continuous function (CYBENKO, 1989; MACHADO, 2012).

In order to perform the supervised learning of the MLPs, the *Back-Propagation* algorithm is frequently used. During the training procedure, the parameters (degrees of freedom or synaptic weights and *bias*) are adjusted aiming to minimize a cost function which allows assessing the quality of the solution (MACHADO, 2012). This procedure is an iterative process in which the updates of the synaptic weight connecting neuron k to neuron j are calculated according to the following equation

$$w_{kj}(n+1) = w_{kj}(n) + \Delta w_{kj}(n) \quad (2.29)$$

where w is the synaptic weight, n is the update step and Δw is the correction applied to the synaptic weight (MACHADO, 2012). The correction $\Delta w_{kj}(n)$ is defined by (HAYKIN, 2009)

$$\Delta w_{kj}(n) = \eta \delta_j(n) y_k(n), \quad (2.30)$$

where η is the learning-rate, the term $\delta_j(n)$ is the local gradient and its value depends on whether the neuron j is located in the output layer or in the hidden layer (HAYKIN, 2009):

- 1 If the neuron j is located in the output layer, $\delta_j(n)$ is equal to the product of the derivative $\Phi_j'(v_j(n))$ and the error signal $e_j(n)$, which is defined by the difference between the NN output $y_j(n)$ and the desired output d_j ;
- 2 If the neuron j is located in a hidden layer, $\delta_j(n)$ is equal to the product of the associated derivative $\Phi_j'(v_j(n))$ and the weighted sum of the δ values computed for the neurons in the next hidden or output layer that are connected to neuron j .

As presented previously, this work proposes the use of NN for TEC modelling in the Brazilian region. The detailed information about the investigated NN models is presented in the next section.

Chapter 3

Results and analysis

This work presents different studies performed in order to investigate the use of Artificial Neural Networks on v TEC estimation in the Brazilian region. These distinct studies consider different input parameters for the network, distinct v TEC calibration techniques, different periods and, in some cases, different stations. The studies will be, henceforward, referred to as Case study 1, 2 and 3. In the first two case studies, the spatial performance of the ANN is assessed. The third case study presents an extended analysis of the ANN model, including its capability to perform short-term forecasting and its comparison with the GIMs.

To perform the analysis, data from some stations of the Brazilian Network for Continuous Monitoring of the GNSS (RMBC) were used on TEC calibration and from the first GLONASS network for research and development (GLONASS R&D network) installed in Latin America were used in this work.

3.1 Preliminaries

This preliminary section presents to the reader some common information for the three Case studies shown in this work. This section includes information about the types of input parameters, employed TEC calibration techniques and the adopted indexes to assess the network performance.

3.1.1 NN model input parameters

TEC magnitude is influenced by solar activity, geographic location, Earth magnetic field and presents diurnal and seasonal variations (HOFMANN-WELLENHOF; LICHTENEGGER; WASLE, 2008; MONICO, 2008). Therefore, the chosen input parameters for the cases under investigation presented in this work can include the latitude, longitude, solar radio flux ($F_{10.7}$), time of the day, day of year (doy), Dst-index and Kp-index. The set of input parameters for each case study will be presented in its corresponding section.

Geographic location

Geographic location has a strong influence in the ionosphere electron density variation (BHUYAN; BORAH, 2007; MONICO, 2008). Although a gross generalization, equatorial regions are characterized by high levels of electron density, mid-latitude regions are considered relatively free of ionospheric anomalies and polar regions of the Earth are very unpredictable (WEBSTER, 1993).

Solar radio flux

The solar radio flux ($F_{10.7}$) index is a determination of the strength of the solar radio emission in a 100 MHz-wide band centered on a 2800 MHz, corresponding to a 10.7 cm wavelength. It is one of the most common indexes used to interpret solar activity. Although it is commonly referred to as a flux, this index is a flux density (TAPPING, 2013). Several fields can be benefited by the research and prediction of $F_{10.7}$, such as the safety of spacecraft, the safety of space and air travel, and others (ZHAO; HAN, 2008).

Solar radio flux data used in this work were obtained from Laboratory for Atmospheric and Space Physics - University of Colorado Boulder (LASP) database, available on-line ¹.

Seasonal variation

Season is another factor that influences electron density variation due to changes in the zenith angle of the Sun and in the radiation flux intensity. During equinoxes, ionospheric effects are greater, whereas during the solstices the ionosphere effects are smaller (HOFMANN-WELLENHOF; LICHTHENEGGER; WASLE, 2008; MONICO, 2008).

Different periods are considered in this work, the first (from day 154 to 163) is a ten-days period at the end of Brazilian autumn, and the second period (from day 282 to 291) is a ten-days period on Brazilian spring. A third period, from day 264 to 268, which includes the September equinox of 2016 is investigated in the Case study 2.

Dst-index

The Disturbance Storm Time (Dst) index was designed to reveal the global magnetic field reductions during storms. This is one of the most widely used indexes to describe space storm activity. This index is given in nanotesla and it is provided on an hourly basis.

Dst data used in this work were obtained from the World Data Center (WDC) for Geomagnetism archive available on-line ².

¹http://lasp.colorado.edu/lisird/data/noaa_radio_flux/

²<http://wdc.kugi.kyoto-u.ac.jp/dstdir/index.html>

Geomagnetic Kp-index

One way to measure the response of Earth magnetosphere due the solar activity is by analyzing the Kp-index. For each three hours, this index indicates the intensity of magnetic activity. Its scale varies from 0 to 9 and is expressed in the following order 0o 0+ 1- ... 8+ 9- 9o, in a scale of thirds (i.e 0+ corresponds to $0 + 1/3$) (BARTELS, 1957). Therefore, one-day data can be expressed as an eight-length vector, with the first value valid from 00 to 03 UT and so on (BARTELS, 1957).

According to the National Oceanic and Atmospheric Administration (NOAA) Geomagnetic Storm scale available on-line³, a Kp-index value greater than 5o indicates geomagnetic storm condition. Kp-index data used in this work were obtained from the GFZ German Research Centre for Geosciences archive available on-line⁴.

3.1.2 Calibration techniques

The vTEC values employed in this work were obtained by using two different calibration techniques. For the Cases of study 1 and 2, the calibrated vTEC values were obtained using the GPS-TEC Analysis Software and for the Case study 3 the ICTP calibration technique was adopted. More information about both methods is presented below.

ICTP calibration technique

Based on the first order approximation described in Section 2.3.4.1, one can obtain an ionospheric observable related to the satellite receiver slant TEC (*sTEC*) from the subtraction of simultaneous observations made at different frequencies. This ionospheric observable can be obtained from either carrier-phase or code-delay measurements (CIRAOLLO et al., 2007). It is important to note that, although carrier-phase measurements observations are much less affected by measurement noise and multipath when compared to code-delay observations, carrier-phase observations present the inconvenient of being biased by unknown ambiguities (CIRAOLLO et al., 2007; MANUCCI et al., 1998).

In order to reduce ambiguities from the carrier-phase ionospheric observables, the leveling carrier to code algorithm is widely employed. In this procedure, the carrier-phase observable ($\tilde{L}_{I,arc}$) is levelled to code-delay ionospheric observable, as shown in Equation (3.1).

$$\tilde{L}_{I,arc} = sTEC + b_R + b_S + \langle \varepsilon_P \rangle_{arc}, \quad (3.1)$$

where $\langle \varepsilon_P \rangle_{arc}$ is the average of the effect of noise and multipath for code-delay observations in a continuous arc. One can note that the carrier-phase inter-frequency biases (IFBs) are

³<http://www.swpc.noaa.gov/NOAA_scales/>

⁴<<ftp://ftp.gfz-potsdam.de/pub/home/obs/kp-ap/>>

replaced by code-delay receiver and satellite biases b_R and b_S . Note that there is no ambiguity term in code-delay observations and the noise and multipath on carrier-phase measurements has been neglected since it is around 100 times less than for code-delay observations (BRAASCH, 1996; CESARONI et al., 2015; CIRAOLLO et al., 2007).

Although it is usually assumed that $\langle \varepsilon_P \rangle_{arc}$ can be disregarded (CESARONI et al., 2015; BRAASCH, 1996), an experiment carried out by Ciruolo et. al 2007 using co-located stations suggests maintaining the term $\langle \varepsilon_P \rangle_{arc}$ trying to estimate bias by arc. Therefore, Equation (3.1) can be written as

$$\tilde{L}_{I,arc} = sTEC + \beta_{arc}, \quad (3.2)$$

which is the basic relation used to TEC calibration, where β_{arc} corresponds to a constant offset to be determined for each arc of observations for a given receiver-satellite pair. β_{arc} represents the contribution of any non-zero error over an arc of observations, e.g. the multipath, and also represents the contribution of receiver and satellite biases, b_R and b_S respectively (CESARONI et al., 2015). In order to define the mapping function from slant TEC to vertical TEC, a two-dimensional thin shell model at 350 km is used. For some stations it is possible to perform TEC calibration from the described method by using an on-line tool provided by ICTP ⁵. The calibrated vTEC values used in this work were obtained in collaboration with the ICTP.

Gopi Seemala calibration technique

The calibrated vTEC values derived from this technique were obtained by using the software GPS-TEC Analysis Application v 2.9.3 available on-line ⁶ and developed by Dr. Gopi Seemala.

The software was used to process the RINEX (Receiver Independent Exchange Format) files for each station and then used to calibrate the vTEC values used in the Case studies 1 and 2. Like ICTP calibration method presented previously, this software also considers a two-dimensional thin shell model at 350 km to define the mapping function from slant TEC to vertical TEC (SEEMALA, 2017).

3.1.3 The ANN models

The MLP ANN with two layers of neurons is used in all case studies presented in this work. Moreover, only one hidden layer was used since it has been shown that the inclusion of additional hidden layers do not lead to much difference in the result accuracy (HABARULEMA et al., 2009).

Figure 3.1 illustrates a ANN structure containing latitude, longitude, doy, time of the day,

⁵<http://t-ict4d.ictp.it/nequick2/gps-tec-calibration-on-line>

⁶<http://seemala.blogspot.com.br/>

Dst-index and solar radio flux as input parameters. Since the problem under investigation is a non-linear process, the activation functions applied to the hidden neurons can be the hyperbolic tangent sigmoid function presented in Equation (2.25) or the logistic sigmoid function described in Equation (2.24) depending on the case study.

The number of artificial neurons in the hidden layer also varies depending on the case study. For the Cases studies 1 and 3, this amount is fixed and it is equal to ten. For the Case study 2, the number of neurons in the hidden layers changes from 1 to 20 in order to obtain the suitable structure as will be presented in Section 3.2. After this procedure, the selected structure is used to estimate the vTEC in the regions and period of interest.

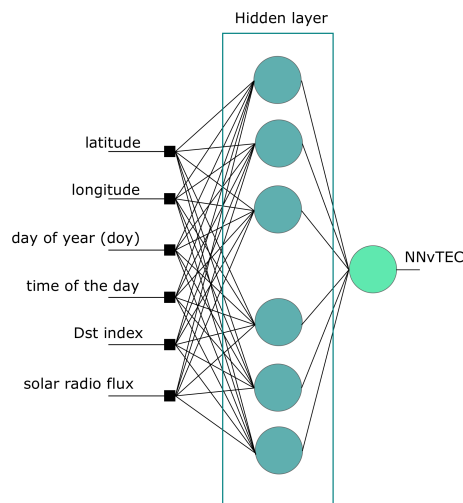


Figure 3.1: Schematic representation of one ANN structure.

Training procedure

In a training process, the ANN provides an estimated output vTEC value (NNvTEC) for a known input vector. This NNvTEC is then compared with the expected output value (calibrated vTEC), and the error for each training pattern is sent back to the hidden layers by the back-propagation algorithm for the update of the ANN synaptic weights. The Mean Square Error (MSE) for training patterns are monitored and training procedure continues as long as the MSE decreases. If the MSE continues decreasing, but the maximum number of iterations (equals to 1,000) is reached, the training is stopped.

3.1.4 NN performance evaluation

For each case study, the assessment of the proposed ANN model performance is done in terms of average error, standard deviation of the error, average relative correction, root mean squared error (RMSE) and the normalized root mean squared error (NRMSE).

The average error (α) is defined as the mean value of the difference between vTEC estimated using the ANN model ($NNvTEC$) and calibrated $vTEC$, as shown in Equation (3.3).

The average relative error (ϵ) is represented in Equation (3.4) (LEANDRO; SANTOS, 2007; HABARULEMA; MCKINNELL; CILLIERS, 2007).

$$\alpha = \frac{1}{N} \sum_{i=1}^N NNvTEC_i - vTEC_i, \quad (3.3)$$

$$\epsilon = \frac{1}{N} \sum_{i=1}^N \frac{|NNvTEC_i - vTEC_i|}{vTEC_i} \times 100, \quad (3.4)$$

where $NNvTEC$ and $vTEC$ are $vTEC$ values obtained by the ANN model and the calibration technique, respectively. The difference $(100 - \epsilon)$ gives the percentual average relative correction, indicating the approximated $vTEC$ prediction accuracy for the proposed ANN model, that is, an average relative error of approximately 10% means that the ANN can predict about 90% of the $vTEC$ in average (LEANDRO; SANTOS, 2007; HABARULEMA; MCKINNELL; CILLIERS, 2007).

The RMSE and NRMSE measure the error with respect to the data used to assess the model quality. These values are defined by Equations (3.5) and (3.6) (HABARULEMA; MCKINNELL; CILLIERS, 2007; MARTINS; NEPOMUCENO; BARROSO, 2013):

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (vTEC_i - NNvTEC_i)^2}. \quad (3.5)$$

$$NRMSE = \frac{\sqrt{\sum_{i=1}^N (vTEC_i - NNvTEC_i)^2}}{\sqrt{\sum_{i=1}^N (vTEC_i - \overline{vTEC})^2}}, \quad (3.6)$$

where $NNvTEC$ is ANN estimative of $vTEC$, whose average is \overline{vTEC} with N corresponding to the number of observations. When NRMSE index is less than one, the models are considered representative, which means that the error is, on average, less than the error given by the mean of the time series (MARTINS; NEPOMUCENO; BARROSO, 2013).

3.2 Case study 1

In this Case study the performance of different MLP ANN structures on the estimation of $vTEC$ are assessed. The ingestion of different parameters which measure solar and geomagnetic activity levels is investigated, including Kp-index, solar radio flux (F10.7) and Dst-index. Furthermore, different activation functions and different amount of neurons are adopted in order to determine the more suitable structure to estimate $vTEC$ in the analyzed period.

3.2.1 Training the NN

In total, 120 MLP structures with different input parameters, different amount of neurons in the hidden layer, different activation functions were tested in order to obtain the more suitable structure to be employed in vTEC for this Case study. The input-output pairs used to train the network consist of an input vector, which varies depending on the structure, and the output is the calibrated vTEC obtained from the GPS-TEC Analysis Software.

Table 3.1 presents the different structures employed in this Case study. The input parameters latitude, longitude, doy and time of the day are not shown in the Table 3.1, since these parameters are adopted in all structures. For each one of the listed structures, the amount of neuron in the hidden layer varies from 1 to 20.

Table 3.1: NN structures under investigation - Case study 1.

Structure	Input Parameters	Activation Function
A	Kp-index, $F_{10.7}$	<i>logsig</i>
B	Dst-index, $F_{10.7}$	
C	Dst-index	
D	Kp-index, $F_{10.7}$	<i>tansig</i>
E	Dst-index, $F_{10.7}$	
F	Dst-index	

The activation functions described in Table 3.1 are obtained from Equations (2.24) and (2.25).

As presented previously, the purpose of this investigation is to select one structure among the structures listed in Table 3.1 to obtain a vTEC model in the South and Central-West regions of Brazil. In order to obtain the most suitable MLP structure, the data from the South region were used. Several structures were tested and the RMSE were assessed. Data from stations RSAL, RSCL, RSPE, SCLA from doy 264 to 268 were presented to the network during the training procedure. The location of the stations are presented in Table 3.2

Table 3.2: Summarized information of the stations under investigation in the South region - Case study 1.

Data Network	Station Name	City in Brazil	Lat. (°S)	Lon. (°W)	Type of Use (vTEC Data)
RBMC	RSPE	Pelotas (RS)	31.802	52.418	NN training
	RSCL	Cerro Largo (RS)	28.141	54.755	NN training
	RSAL	Alegrete (RS)	29.789	55.769	NN training
	SCLA	Lages (SC)	27.792	50.304	NN training
	SMAR	Santa Maria (RS)	29.718	53.716	NN testing
	POAL	Porto Alegre (RS)	30.659	51.120	NN testing

After the training procedure, the different structures were used to estimate vTEC on SMAR station during the same period of training. It is important to note that no data from the

test station (SMAR) were presented to the network during the training procedure. Figure 3.2 shows the RMSE obtained for each structure.

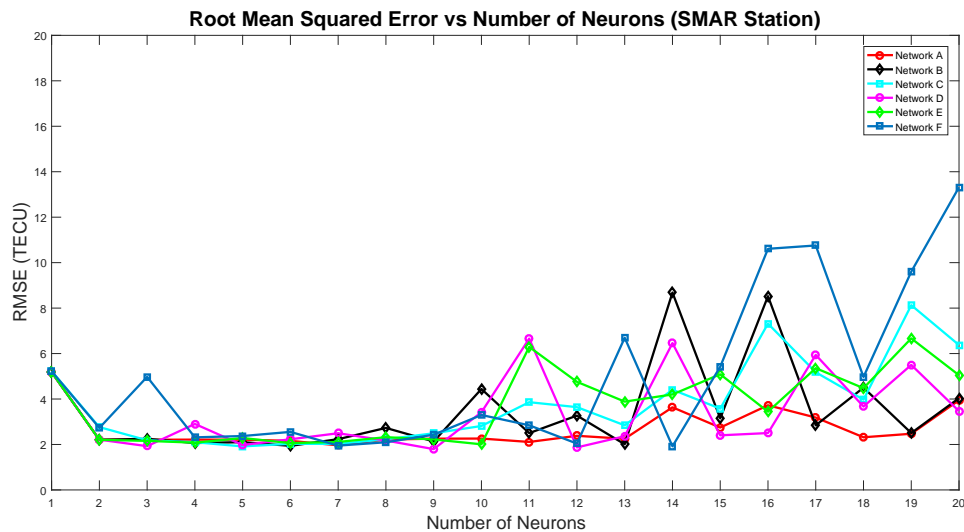


Figure 3.2: RMSE values for each NN structure tested - Case study 1.

By analyzing the Figure 3.2, one can note that the structure D presented the best RMSE when compared to the other structures. The input parameters of this structure are latitude, longitude, doy, time of the day, Kp-index, and solar radio flux (F10.7). It has 9 neurons on the hidden layer and the adopted activation function is the *tansig* described by Equation (2.25). The RMSE obtained using this structure is 1.795 TECU.

3.2.2 Results and discussion

South region

The structure D was employed on v TEC estimation on POAL and SMAR stations during the September equinox of 2016 (doy 266). Whereas the SMAR station was chosen because it is located inside the region delimited by the training stations, the POAL station was selected to test the spatial extrapolation performance of the network, since this station is located out of the region limited by the training stations.

Figures 3.3(a) and 3.3(b) present the v TEC profile estimated by the NN for SMAR and POAL stations, respectively. As shown in Table 3.4, the RMSE obtained was 1.187 and 1.263 TECU for SMAR and POAL, respectively. One can note that the station located inside the region delimited by the training stations, i.e SMAR, presented lowest RMSE.

Central-West region

Aiming to assess the performance of the structure proposed in Section 3.2.1 in different latitude regions, the same estimation procedure of v TEC profile used in the South region was

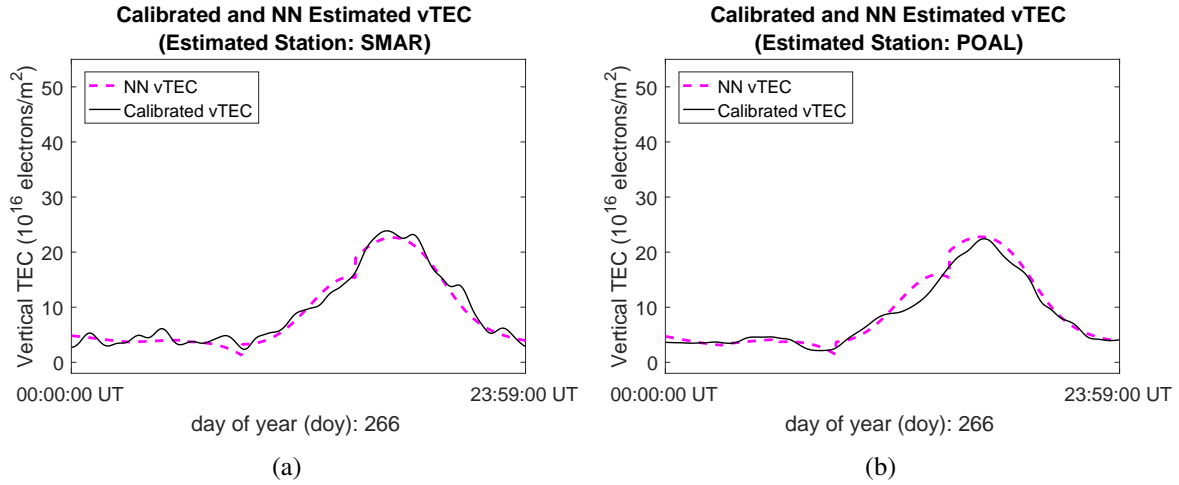


Figure 3.3: Calibrated vTEC and NN vTEC on the September equinox of 2016 for the South region - Case study 1. (a) SMAR station, (b) POAL station.

employed in the Central-West region. In this case, the stations BRAJ, GOJA, GOUR and MGUB were adopted as training stations and the stations GOGY and MTNX were chosen to be the testing stations. The GOGY station was selected because it is located inside the area delimited by the training stations, whereas the MTNX was selected because it is located out of this area. Table 3.3 presents the geographic location of the referred stations.

Table 3.3: Summarized information of the stations under investigation in the Central-West region - Case study 1.

Data Network	Station Name	City in Brazil	Lat. (°S)	Lon. (°W)	Type of Use (vTEC Data)
GLONASS R&D	BRAJ	Brasília (DF)	15.772	47.866	NN training
	GOJA	Jataí (GO)	17.883	51.726	NN training
RBMC	GOUR	Uruaçú (GO)	14.509	49.144	NN training
	MGUB	Uberlândia (MG)	18.919	48.257	NN training
	MTNX	N. Xavantina (MT)	14.697	52.349	NN testing
	GOGY	Goiânia (GO)	16.664	49.255	NN testing

Figures 3.4(a) and 3.4(b) present the estimated vTEC profile for the September equinox of 2016 for the stations GOGY and MTNX, respectively. Also in this case, one can verify that the better results were obtained in the station located inside the area delimited by the training stations. The RMSE values obtained for GOGY and MTNX stations were equal to 1.791 and 3.370, respectively. The results obtained in this Case study are summarized in Table 3.4.

3.3 Case study 2

In this Case study a MLP ANN is used to estimate the vTEC in the Central-West region of Brazil, more specifically on GOGY station. The hidden layer of the MLP consists of

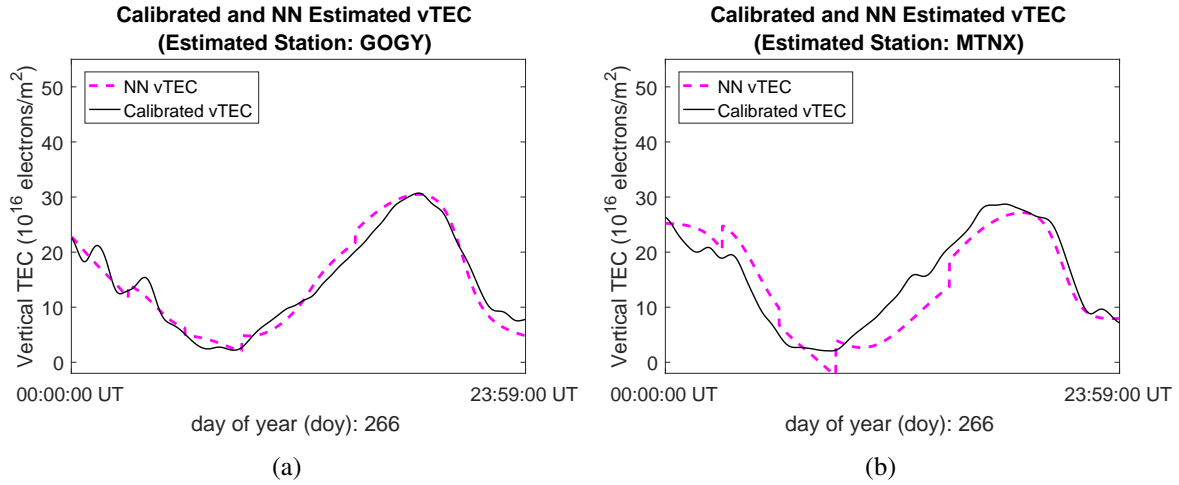


Figure 3.4: Calibrated vTEC and NN vTEC on the September equinox of 2016 for the Central-West region - Case study 1. (a) GOGY station. (b) MTNX station.

Table 3.4: Spatial performance of the NN model - Case Study 1.

Station (region)	Analysis period (doy)	NRMSE	RMSE (TECU)	α (TECU)	Std. dev. of the error (TECU)	Rel. correction (%)
POAL (SO)	266	0.2057	1.253	0.544	1.129	89.15
SMAR (SO)	266	0.1769	1.187	-0.365	1.130	89.67
GOGY (CW)	266	0.2126	1.791	0.062	1.791	89.59
MTNX (CW)	266	0.4311	3.730	-1.049	3.580	79.94

10 artificial neurons and the output layer has only one artificial neuron. For each set of input data (latitude, longitude, doy, time of the day, Kp-index, and solar radio flux), the network provides an estimated vertical TEC ($NNvTEC$) value. Since the problem under investigation is a non-linear process, the activation function applied to the hidden neurons is the non-linear sigmoid function shown in Equation (2.24), with a equals to 1. A linear function was considered for the output layer.

Table 3.5 presents location information of the stations used in this study. A map to illustrate the station distribution is presented in Figure 3.5.

Table 3.5: Summarized information of the stations under investigation - Case study 2.

Data Network	Station Name	City in Brazil	Lat. ($^{\circ}$ S)	Lon. ($^{\circ}$ W)	Type of Use (vTEC Data)
GLONASS R&D	BRAJ	Brasília (DF)	15.772	47.866	NN training
	BRAZ	Brasília (DF)	15.947	47.878	NN training
	GOUR	Uruaçu (GO)	14.509	49.144	NN training
RBMC	MTBA	Barra do Garças (MT)	15.876	52.265	NN training
	MGUB	Uberlândia (MG)	18.919	48.257	NN training
	GOGY	Goiânia (GO)	16.664	49.255	NN testing

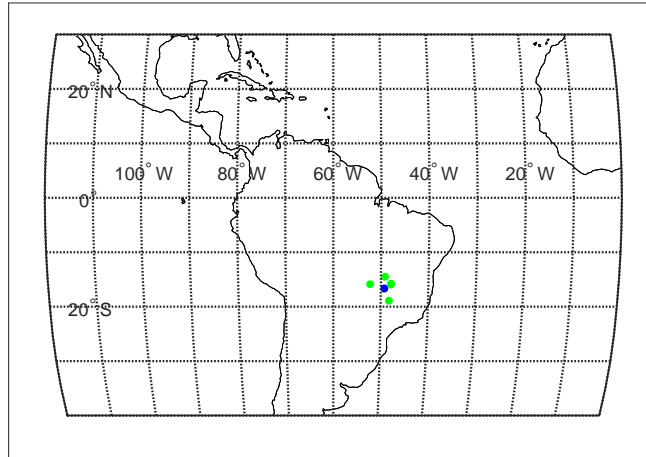


Figure 3.5: Positions in the world map of the stations under investigation - Case study 2. Green and blue dots represent training and testing station, respectively.

3.3.1 Training the NN

In this scenario, $vTEC$ data from stations GOUR, MTBA, MGUB and BRAJ were used in the training process. The pair input-output used to train the network consists of an input vector containing the latitude, longitude, doy, time of the day, Kp-index, and solar radio flux and the output is the $vTEC$ calculated using the GPS-TEC Analysis Application described in Section 3.1.2. Input-output data from doy 154 to 163 were applied randomly to the network during the training process.

Due to the absence of data for some stations in the considered interval, a total of 43200 samples were presented to the network during the learning process. In order to avoid overfitting or overtraining, 85% of the training set was used to the training procedure and 15% was used to validate the model. This procedure is important since the network may lose its generalization ability if overtrained (HAYKIN, 2009). The Levenberg-Maquardt backpropagation algorithm was used to set the ANN weights.

Figure 3.6 presents the performance of the ANN for each epoch during the training and validation procedures. The best validation performance was obtained at the epoch 159, corresponding to a Mean Square Error of 2.750 TECU^2 .

3.3.2 Results and discussions

After the training procedure, the network is then used to estimate the $vTEC$ profile for GOGY station from doy 154 to 163. Figure 3.7 presents the $vTEC$ estimated by the ANN and the $vTEC$ calculated using the GPS-TEC Analysis Application for the GOGY station.

Table 3.6 summarizes some statistics of the results obtained with the proposed ANN

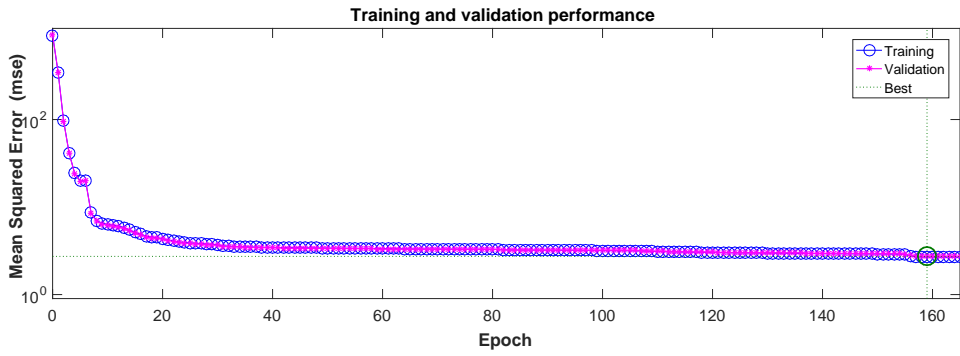


Figure 3.6: Performance of the ANN during the training and validation procedures - Case study 2.

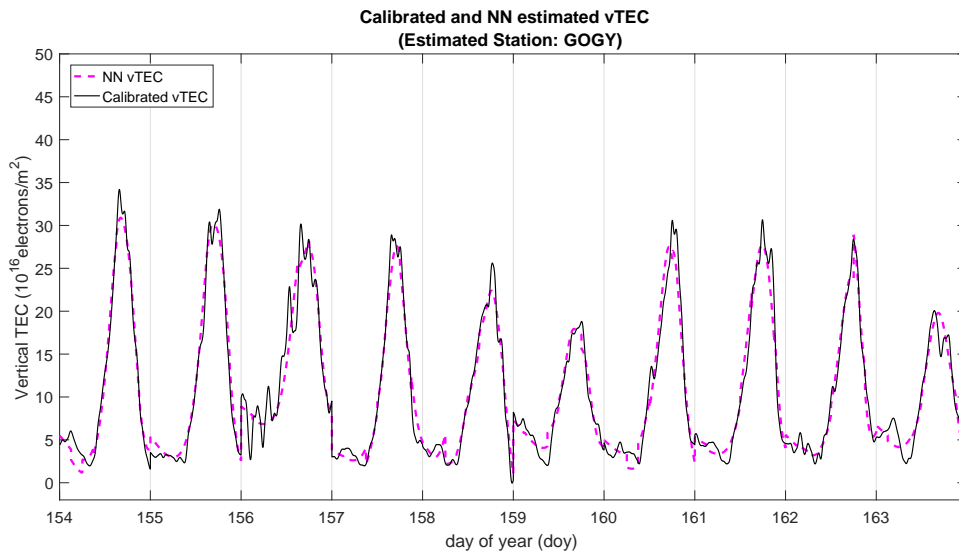


Figure 3.7: Calibrated vTEC and NN vTEC on GOGY station - Case Study 2.

model. As shown in Table 3.6, the average error α of the estimative is -0.016 TECU and the standard deviation of the error is 1.635 TECU. In terms of relative error ϵ , it was found an average value of 11.19%. Alternatively, indicating an agreement in 88.81% of the cases on average.

Table 3.6: Spatial performance of the NN model - Case study 2.

Station (region)	Analysis period (doy)	NRMSE	RMSE (TECU)	α (TECU)	Std. dev. of the error (TECU)	Avg. rel. correction (%)
GOGY (CW)	154 to 163	0.1948	1.635	-0.016	1.635	88.81

3.4 Case study 3

In this Case study, the vTEC values on three distinct Brazilian regions (Northeast, Central-West and South) are estimated. Data from two ten-days periods (from doy 154 to 163 and from 282 to 291) are used to train the ANN. By using this dataset, three distinct

analyses have been carried out in order to assess time-varying and spatial performance of the ANN model.

At the spatial performance analysis, for each region, a set of stations is chosen to provide training data to the ANN and after the training procedure the ANN is used to estimate vTEC values for the test station which data were not presented to the ANN during the training process. An analysis is done by comparing, for each testing station, the estimated NNvTEC delivered by the ANN and the reference calibrated vTEC. As a second analysis, the NN ability to forecast one day after the time interval (doy 292) based on information of the second period of investigation is also assessed in order to verify the feasibility of using a low amount of data for short-term forecasting. In a third analysis, the spatial performance of the ANN model is assessed and compared against CODE Global Ionospheric Maps during the geomagnetic storm registered on 13th and 14th October 2016.

3.4.1 Training the NN

Aiming to evaluate the performance of the proposed ANN model, vTEC values of North-east, Central-west and South regions of Brazil are investigated. To accomplish this task, for each region, a set of stations is selected to provide training data to the ANN and one station (out of the training set) is used to test the network and measure its performance. It is important to notice that no data from the test station are presented to the ANN during the training process.

Known input and output data are required to train the network. The input vector consists of latitude, longitude, doy, time of the day, Kp-index, and solar radio flux, and the output is the calibrated vTEC using the ICTP calibration technique described in Section 3.1.2.

The map shown in Figure 3.8 presents the positions of the stations under analysis for each region. The training and test stations are represented by the green and blue dots, respectively.

Data from two different periods in 2016 are used in the training procedure. The first and second periods will be referred to as *Period A* and *Period B*, respectively. Period A corresponds to doy 154 to 163 (from June 2nd to June 11th, 2016). Average solar radio flux (10.7 cm) is 84.77 sfu with a standard deviation of 3.59 sfu, for this interval. Period B corresponds to doy from 282 to 291 (from October 8th to October 17th, 2016) and presents an average solar radio flux (10.7 cm) of 94.33 sfu with a standard deviation of 10.06 sfu.

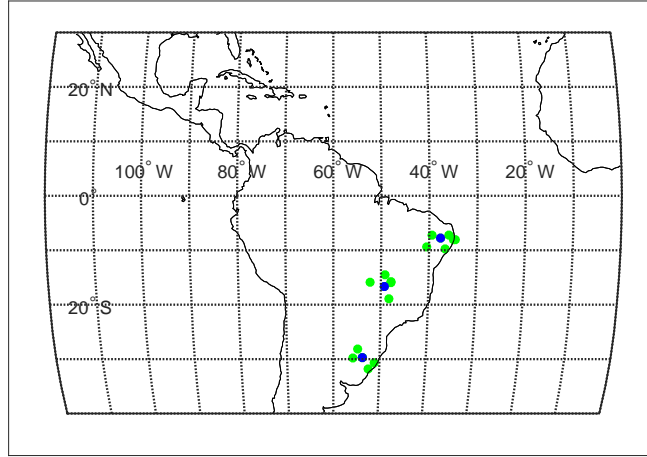


Figure 3.8: Positions in the world map of the stations under investigation - Case study 3. Green and blue dots represent training and testing stations, respectively.

3.4.2 Results and discussion

3.4.2.1 Assessment of NN spatial performance

Northeast region

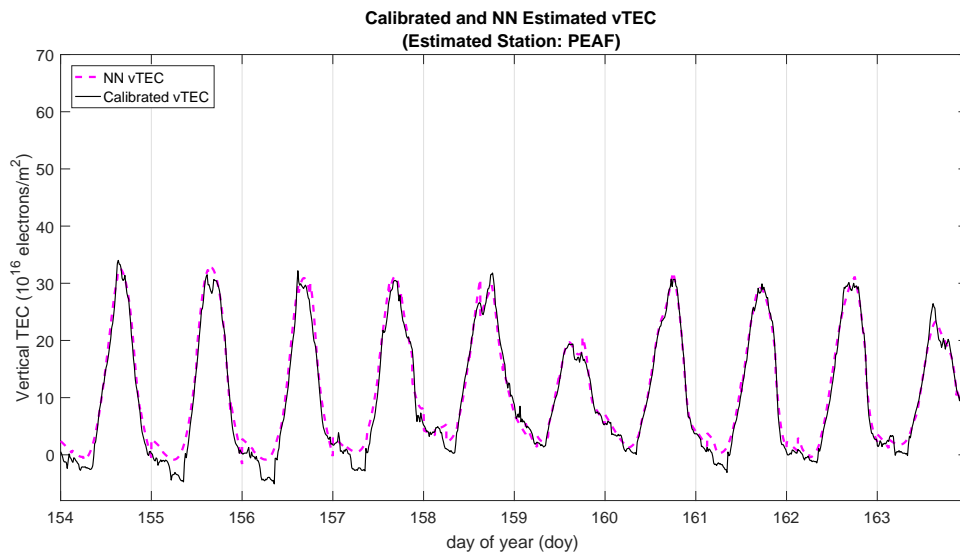
In order to train the ANN for this region, data derived from six stations were used. Table 3.7 presents information about geographic location of the stations exploited in Brazilian Northeast region.

Table 3.7: Summarized information of the stations under investigation in the Northeast region - Case study 3.

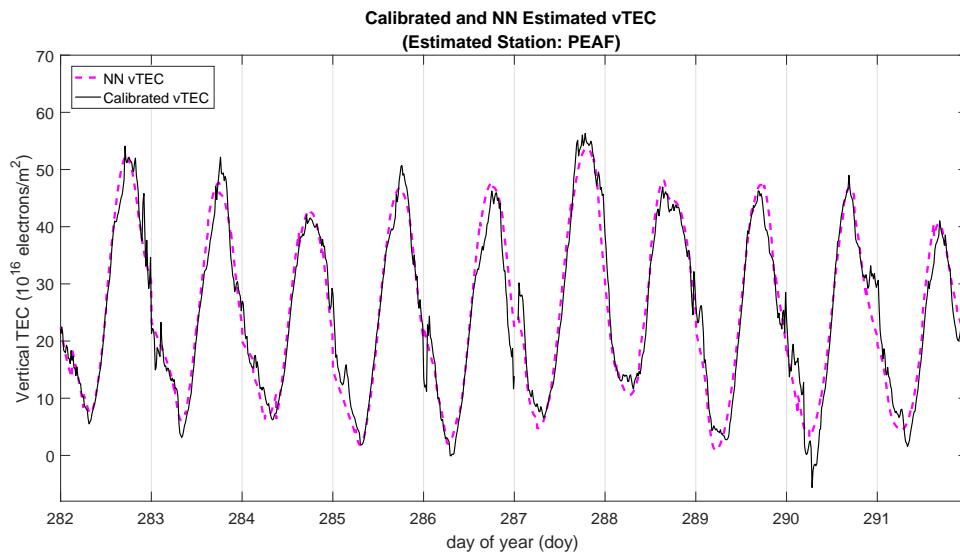
Data Network	Station Name	City in Brazil	Lat. (deg S)	Lon. (deg W)	Type of Use (vTEC Data)
GLONASS R&D	RSFJ	Recife (PE)	8.059	34.543	NN training
	ALAR	Arapiraca (AL)	9.749	36.654	NN training
RBMC	CRAT	Crato (CE)	7.228	39.416	NN training
	PBCG	Campina Grande (PB)	7.213	35.908	NN training
	PEPE	Petrolina (PE)	9.348	40.507	NN training
	RECF	Recife (PE)	8.050	34.952	NN training
	PEAF	Afог. da Ingazeira (PE)	7.714	37.632	NN testing

After the training procedure, the ANN model was used to estimate the vTEC on PEAF station during Period A and Period B. Figure 3.9 presents the result obtained by using the proposed ANN model for both periods. In this case, data are provided and plotted on a 15 seconds scale.

One can note that peak vTEC values obtained during Period A are lower than values obtained on Period B. This result indicates the influence of solar activity in vTEC value,



(a)



(b)

Figure 3.9: Calibrated vTEC and NN vTEC on the PEAf station - Case Study 3. (a) Period A. (b) Period B.

since this parameter is greater on Period B. Furthermore, during day 287 and 288, it is possible to verify on Figure 3.9(b) an increasing of vTEC values when compared with the other days of Period B, indicating some correspondence with the storm occurred on 13th and 14th October 2016. This event and its influence on stations under analysis will be presented with more details in Section 3.4.2.3.

Considering Period A, the Table 3.10 shows that the ANN model estimated the vTEC with an average relative correction of approximately 86.6% and an average error of 0.942 TECU. The ANN model estimation for Period B provided an average error of approximately 0.214 TECU and an average relative correction of 89.4%. The NRMSE obtained for the periods A and B are 0.170 and 0.228, respectively.

Central-west region

Data from stations BRAJ, BRAZ, GOUR, MTBA, MGUB were used to train the network in this region. For both periods, station GOGY, located in Goiânia-GO, was used to test the NN. Table 3.8 indicates geographic position of each station considered in this region.

Table 3.8: Summarized information of the stations under investigation in the Central-West region - Case study 3.

Data Network	Station Name	City in Brazil	Lat. (°S)	Lon. (°W)	Type of Use (vTEC Data)
GLONASS R&D	BRAJ	Brasília (DF)	15.772	47.866	NN training
	BRAZ	Brasília (DF)	15.947	47.878	NN training
RBMC	GOUR	Uruaçu (GO)	14.509	49.144	NN training
	MTBA	Barra do Garças (MT)	15.876	52.265	NN training
	MGUB	Uberlândia (MG)	18.919	48.257	NN training
	GOGY	Goiânia (GO)	16.664	49.255	NN testing

Following the same ANN training and testing procedures described previously, but using stations listed in Table 3.8, the results shown in Figure 3.10 were obtained. One can note that vTEC on GOGY station presented similar behavior on day 287 and 288 when compared with PEA station, but presenting peak values slightly greater during these two days.

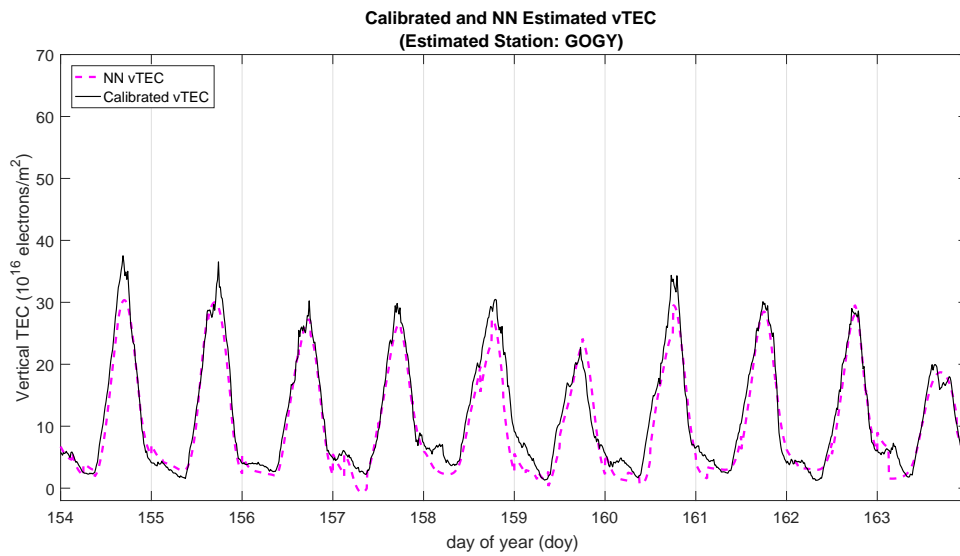
By observing vTEC values shown in Figure 3.10, one can note that the ANN model underestimates vTEC for the most of the analyzed period. This fact can be confirmed by looking at mean values of calibrated vTEC and NNvTEC presented in the Table 3.10. The average error value of vTEC predicted by the ANN model was 10.39 TECU against 11.70 TECU obtained from the calibrated vTEC, leading to a mean error α of about -1.310 TECU.

By analyzing the Period A, the NN model provided an average relative correction of approximately 82.9%. For Period B, the average relative correction is about 88.4%. The ANN vTEC estimative for this station presented a NRMSE of 0.298 and 0.262 for periods A and B, respectively.

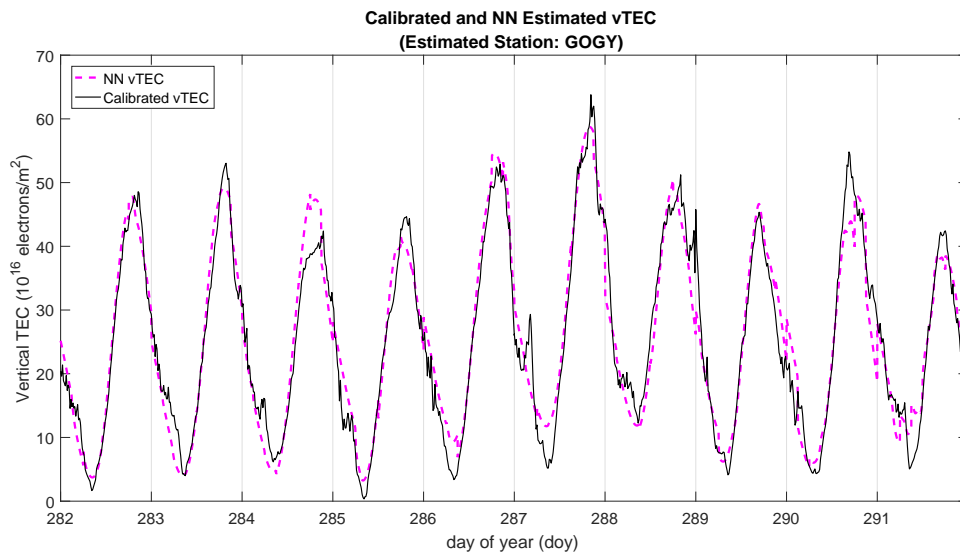
South region

Data from stations RSPE, RSCL, RSAL and POAL were used in the training procedure for South region in Brazil. Table 3.9 indicates the geographic position of each station under study. Figure 3.11 presents the results with a temporal resolution of 30 seconds due the data interval provided by SMBJ station. By comparing Figures 3.9 and 3.10 with Figure 3.11, it is possible to verify that the vTEC peak obtained on PEA and GOGY is approximately two times greater than the vTEC peak calculated for SMBJ in Period A. On period B, the difference is smaller.

As shown in Figure 3.11(b), during day 287 and 288, TEC values of this station was increased, as well. Assessing the results with the criteria described previously, for Period



(a)



(b)

Figure 3.10: Calibrated vTEC and NN vTEC on GOGY station - Case study 3. (a) Period A. (b) Period B.

Table 3.9: Summarized information of the stations under investigation in the South region - Case study 3.

Data Network	Station Name	City in Brazil	Lat. (°S)	Lon. (°W)	Type of Use (vTEC Data)
GLONASS R&D	SMBJ	Santa Maria (RS)	29.718	53.719	NN testing
	RSPE	Pelotas (RS)	31.802	52.418	NN training
	RSCL	Cerro Largo (RS)	28.141	54.755	NN training
RBMC	RSAL	Alegrete (RS)	29.789	55.769	NN training
	POAL	Porto Alegre (RS)	30.659	51.120	NN training

A, the ANN model provided an average relative correction of approximately 88.6%. For Period B, the average relative correction obtained is about 88.9%. At this location, the ANN

estimative presented NRMSE of 0.265 for period A and 0.214 for period B, respectively.

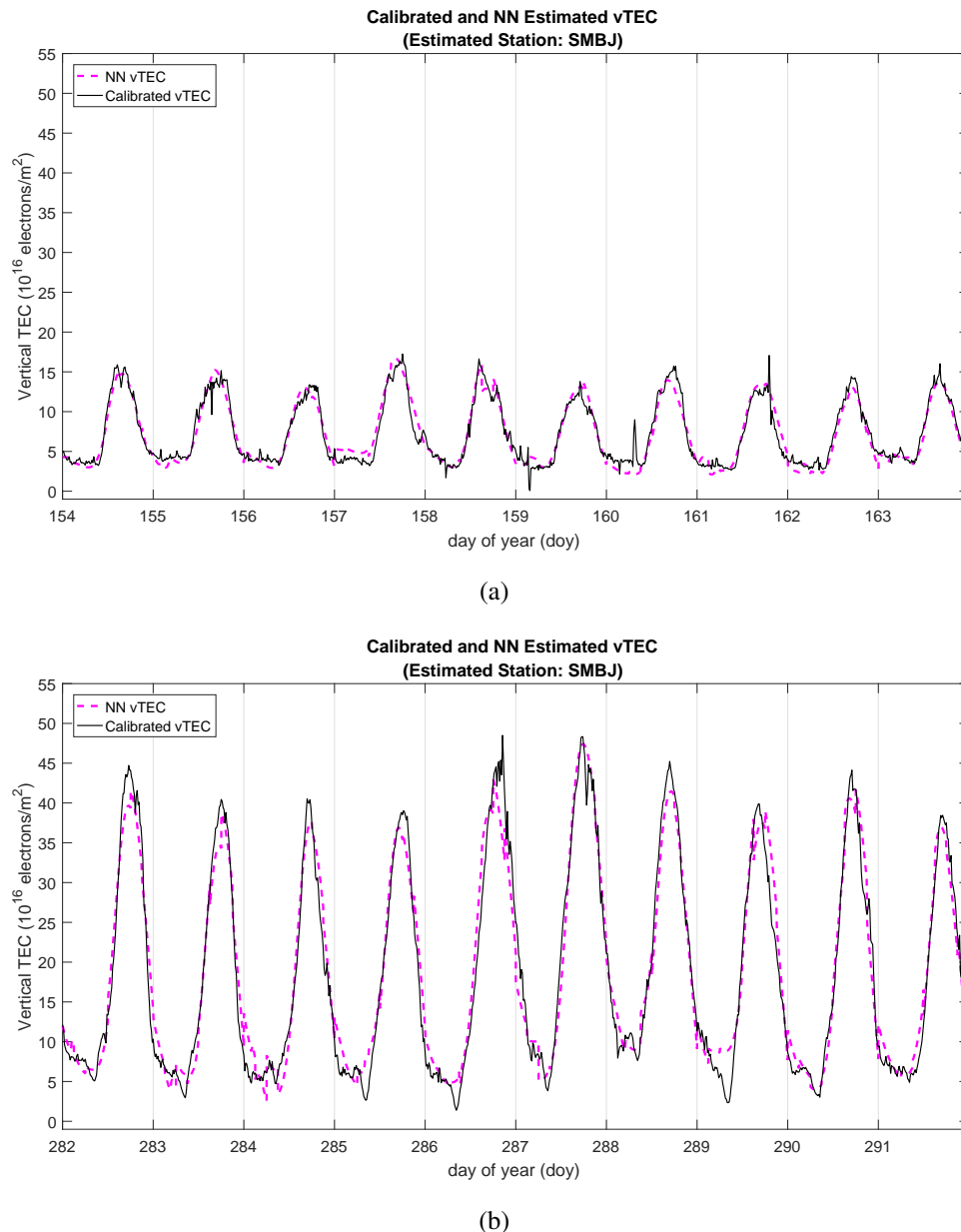


Figure 3.11: Calibrated vTEC and NN vTEC on SMBJ station - Case study 3. (a) Period A. (b) Period B.

Table 3.10 summarizes the results obtained using the proposed ANN model for all three-stations during both periods. It is possible to note that the minimum average relative correction obtained was 82.9% indicating that the proposed NN model has provided good estimative for the observed stations in the considered periods.

3.4.2.2 Assessment of short-term forecasting performance

In order to assess the ANN short-term forecasting ability, training stations data from Period B (day 282 to 291) were presented to the ANN, and then, the ANN was used to

Table 3.10: Spatial performance of the NN model - Case study 3.

Station (region)	Analysis period	NRMSE	RMSE (TECU)	α (TECU)	Std. dev. of the error (TECU)	Rel. correction (%)
PEAF (NE)	A	0.170	1.762	0.942	1.488	86.6
	B	0.228	3.401	0.214	3.394	89.4
GOGY (CW)	A	0.298	2.615	-1.310	2.267	82.9
	B	0.261	3.768	0.289	3.757	88.4
SMBJ (SO)	A	0.265	1.073	0.078	1.070	88.6
	B	0.209	2.663	0.239	2.653	88.9

forecast vTEC on doy 292. It is important to highlight that no data from doy 292 were presented to the ANN in the training procedure. The results for stations PEAFF and GOGY are presented on a 30s basis and for SMBJ station the results are presented in a 15s basis.

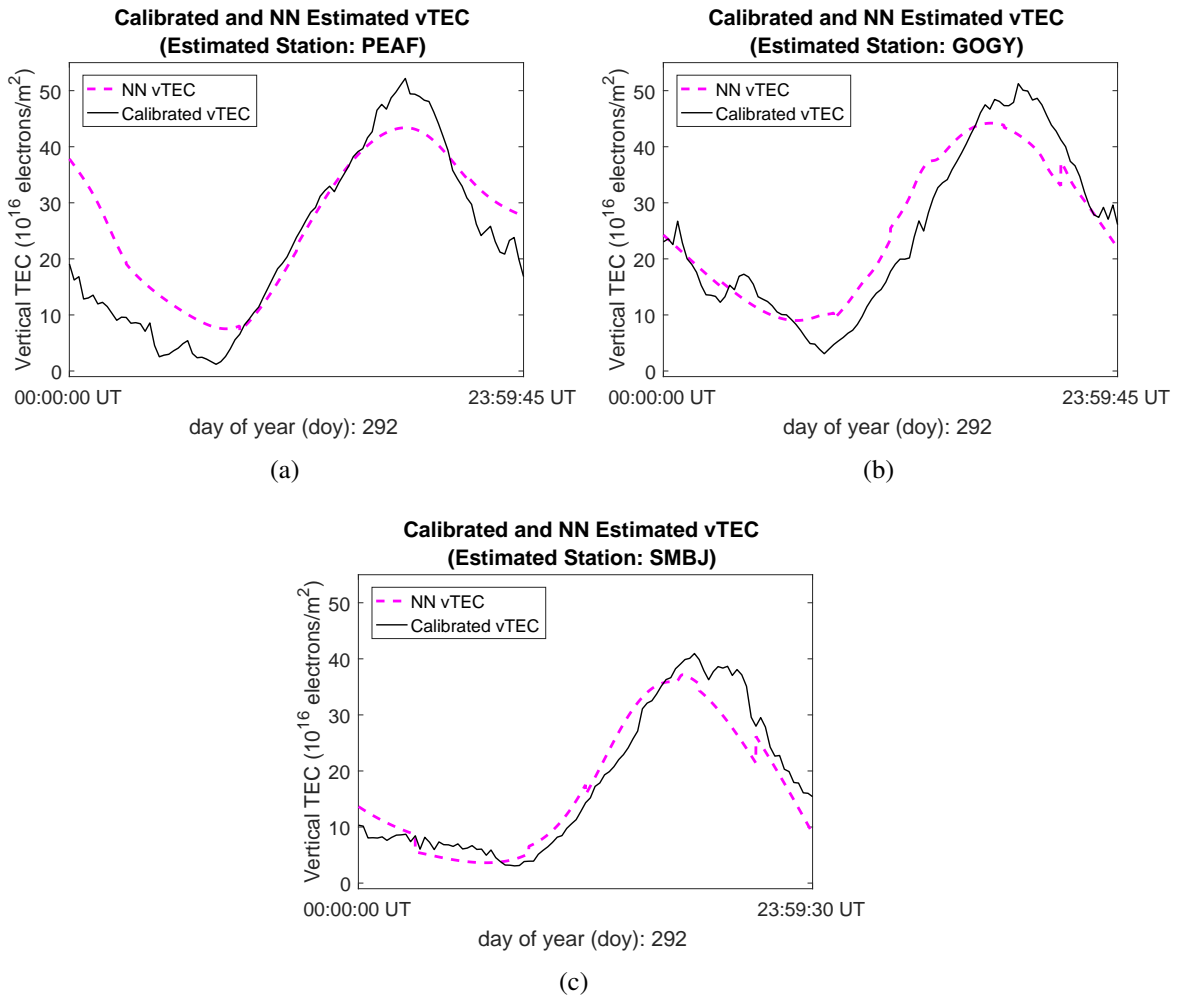


Figure 3.12: Calibrated vTEC and NN vTEC for short-term forecasting of doy 292 - Case study 3. (a) PEAFF station. (b) GOGY station. (c) SMBJ station.

Table 3.11 presents the results obtained for each region under investigation. One can note that the maximum RMSE error obtained for short-time forecasting was found at PEAFF station, corresponding to 7.84 TECU. At this station, the average relative correction observed was 73.5%. Better results were obtained at GOGY and SMBJ stations, with a RMSE of 5.18 and 4.05 TECU, respectively.

Table 3.11: Short-term forecasting performance of the NN model - Case study 3.

Station (region)	NRMSE	RMSE (TECU)	α (TECU)	Std. dev. of the error (TECU)	Avg. rel. correction (%)
PEAF (NE)	0.502	7.84	3.544	6.999	73.3
GOGY (CW)	0.355	5.18	0.577	5.153	82.3
SMBJ (SO)	0.320	4.05	-0.813	3.972	81.6

3.4.2.3 Comparison of calibrated TEC, NN predicted TEC and CODE GIMs

According to the Weekly Space Weather Report 20161020 prepared by The Mexican Space Weather Service (SCiEMEX) and available on-line ⁷, the registered Dst and Kp indexes indicated a geomagnetic storm during 13th and 14th October 2016. Data provided by the World Data Center for Geomagnetism (WDC)⁸ registered Dst indexes reaching -104 and -97 nanotesla for 13th and 14th, respectively. Kp-indexes registered during these days, also indicate geomagnetic storms condition. The registered 3-hours Kp-index are presented in Table 3.12.

Table 3.12: Registered Kp-index values.

Day	Kp-index values							
	00 UT	03 UT	06 UT	09 UT	12 UT	15 UT	18 UT	21 UT
2016 October 13	2+	3o	4+	5-	5-	6+	6+	5o
2016 October 14	6-	5-	5-	3o	2-	2-	1-	1-

According to the NOAA Geomagnetic Storm scale described in Section 3.1.1, the geomagnetic storm conditions were registered on both days. In order to evaluate the performance of the ANN proposed model during this period, a comparison with GIMs provided by CODE is performed.

In the GIMs generated by CODE, the vertical TEC is modeled in a solar-geomagnetic frame by the use of spherical harmonics expansion and piecewise linear functions are used for the representation in the time domain (YU et al., 2015). Daily differential code biases (DCB) for all GPS satellites and ground stations are estimated simultaneously as constant values for each day (YU et al., 2015). GIMs are generated on a daily basis at CODE using data from about 300 GNSS sites of the IGS and other institutions. Final CODE GIMs are provided with a spatial resolution of 2.5° in latitude and 5.0° in longitude with a latency of approximately two weeks. Temporal resolution is 1 hour since GPS week 1815. GIMs used in this work were obtained from the Crustal Dynamics Data Information System (CDDIS) database available on-line ⁹.

In this analysis, the purpose is to compare the vTEC values provided at each hour in the CODE GIMs Ionex files and the vTEC values provided by the ANN in the same epochs. Performance comparison of both methods will be evaluated by RMSE, average Δ TEC and

⁷<http://www.sciesmex.unam.mx/>

⁸<http://wdc.kugi.kyoto-u.ac.jp/dstdir/index.html>

⁹<ftp://cddis.gsfc.nasa.gov/gnss/products/ionex/>

average relative correction considering calibrated vTEC as the reference. Average ΔT_{av} is defined as (HABARULEMA et al., 2009):

$$\Delta T_{av} = \frac{1}{N} \sum_{i=1}^N ((T_{j=1,2})_i - T_i), \quad (3.7)$$

where ΔT_{av} is the average difference between the calibrated GNSS vTEC (T) and NN TEC ($T_{j=1}$) or GIM TEC ($T_{j=2}$), $i = 1, 2, \dots, N$ with N corresponding to the number of observations at a particular time during the interval.

Figure 3.13 presents the calibrated vTEC, NNvTEC and GIM vTEC from doy 286 (0h00 UT) to doy 288 (23h00 UT). Calibrated vTEC and NNvTEC data are not available for PEAf station in 0h00 UT of doy 287, therefore, this observation will not be plotted or considered in the computation of performance.

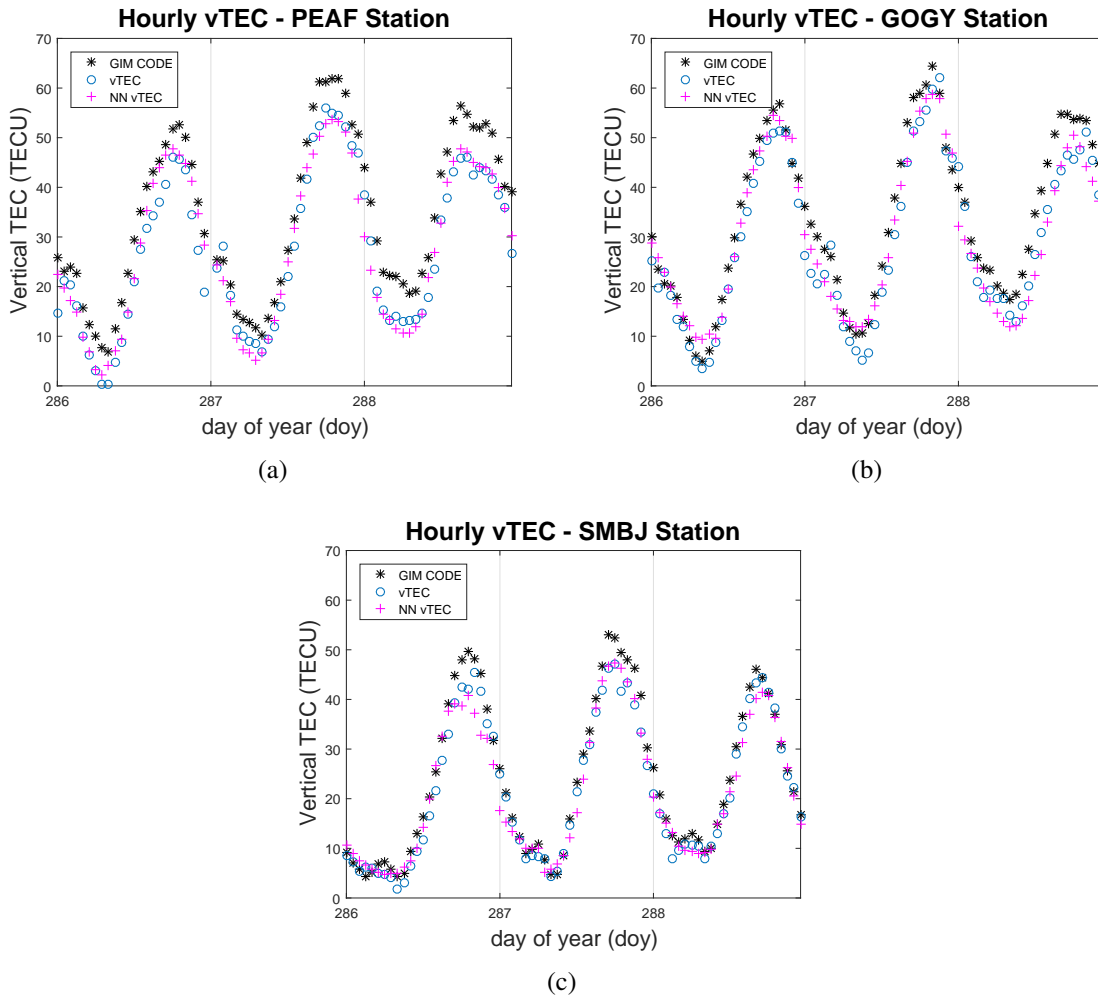


Figure 3.13: Calibrated vTEC, NN vTEC, and GIM derived vTEC from doy 286 to 288 on year 2016 - Case study 3. (a) PEAf station. (b) GOGY station. (c) SMBJ station.

As shown in Figure 3.13, the proposed NN presented better results when compared with CODE GIMs during the period under investigation (considering vTEC provided by the calibration as the reference). In almost all period, GIMs overestimate TEC values in all stations.

Table 3.13 presents the performance of both applied methods.

Table 3.13: Performance evaluation of NN vTEC and GIM vTEC with respect to calibrated TEC from doy 286 to 288.

Station (region)	NRMSE		RMSE (TECU)		ΔT_{av} (TECU)		Std. dev. of the error (TECU)		Avg. rel. correction (%)	
	NN	GIM	NN	GIM	NN	GIM	NN	GIM	NN	GIM
PEAF (NE)	0.226	0.475	3.50	7.34	0.51	6.84	3.48	2.68	90.42	74.49
GOGY (CW)	0.252	0.358	3.98	5.66	0.65	4.56	3.96	3.36	89.00	83.06
SMBJ (SO)	0.206	0.229	2.92	3.24	-0.21	2.35	2.93	2.25	89.85	88.23

By the analysis of Table 3.13, it is possible to note that the ANN model performed better when compared with CODE GIMs for the period. Minimum and maximum average relative error obtained by the ANN was approximately 89.00% and 90.43% against 74.49% and 88.23% for CODE GIMs. Furthermore, the average difference between calibrated vTEC and CODE GIMs confirms the observed over-prediction of TEC values during the period shown in Figure 3.13. Results presented in Table 3.13 indicate that the ANN model presented better performance (assuming the calibrated vTEC as the reference) when compared with CODE GIMs in all stations analyzed in the interval.

Chapter 4

Conclusion

The results obtained by the analysis performed in the cases of study presented in this work indicate that the use of the proposed Artificial Neural Network models for vertical TEC estimation can provide good estimates in the Brazilian region. The studies performed in the case study 1 presented the impact of choosing different parameters to compose the ANN model. For this case study, the use of the Kp-index and the solar radio flux presented better results. Also, it was possible to note that the number of neurons employed in the ANN is an important parameter to be considered. A large number of neurons does not lead necessarily to a performance improvement for the investigated ANNs. For the case study 2, the proposed ANN model has provided an average relative correction of 88.81% and a RMSE value of 1.635 TECU.

Considering the case study 3, one can verify that the proposed model has provided a minimum relative correction value of approximately 82.9% and a maximum RMSE of 3.76 TECU. Periods of higher solar activity presented the worst average absolute errors results. Periods of lower solar activity provided worst relative error results, due to the lower absolute TEC during days with low solar activities. By analyzing the average absolute error of the ANN model it is possible to verify that better results were obtained during the lower solar activity period.

In the forecasting performance analysis presented in the case study 3, the maximum RMSE value obtained was 7.84 TECU at PEAFF station. Best results were obtained for GOGY and SMBJ stations, reaching RMSE values of 5.18 and 4.05 TECU, respectively. The obtained estimation suggests that the proposed ANN can be exploited on the development of a TEC short-term prediction model in Brazil. Future studies will be carried out in order to determine the suitable training interval to obtain minimum error values in short-term forecasting and improve the NN performance in the regions located at lower latitudes. Also, the use of recurrent ANNs will be investigated.

Investigation of performance during the period from DoY 286 to 288 shows that the ANN model provided better results when compared with CODE GIMs, assuming the calibrated

vTEC as the reference. Performance assessed by this work suggests that the proposed ANN model is a promising tool for vTEC estimation in void areas even during events such as the geomagnetic storm registered during 13th and 14th October 2016. It is important to note that work does not intend to undervalue the estimations provided by CODE Global Ionospheric Model, but rather, reinforce its important role in studies and improvements of regional models.

It is worth mentioning that the proposed ANN model does not have the ability to calibrate vTEC by itself, relying on data provided by the calibration technique. Therefore, if vTEC delivered by the calibration procedure present negative values, the ANN can incorporate this results, and as a consequence presents negative vTEC.

The studies presented in this manuscript were the basis for the investigations published in three papers. The list of publications derived from this research can be found in Appendix A. In future works, the investigations on the development of a model that includes the five Brazilian regions can be performed. The applicability in short-term forecasting in a near real-time scenario will also be investigated. In addition, the performance analysis of low-cost single-frequency receivers using the TEC values from the proposed ANN to mitigate the ionospheric range error in positioning systems will also be investigated.

REFERENCES

- AGUIAR, C. R. *Modelo Regional da Ionosfera Mod ion: implementação em tempo real*. Master Thesis (MSc) — Faculdade de Ciências e Tecnologia UNESP, Presidente Prudente, 2005.
- ALKAN, R. M.; KAMMAN, H.; SAHIN, M. GPS, GALILEO and GLONASS Satellite Navigation Systems & GPS Modernization. In: IEEE. *Proceedings of 2nd International Conference on Recent Advances in Space Technologies*. Istanbul, Turkey, 2005. p. 390–394.
- APPLETON, E. V. Wireless studies of the ionosphere. In: INSTITUTION OF ELECTRICAL ENGINEERS. *Journal of the Institution of Electrical Engineers*. 1932. v. 71, no. 430, p. 642–650.
- ARBESSER-RASTBURG, B.; ROGERS, D. Earth–space propagation. In: BARCLAY, L. (Ed.). *Propagation of Radio Waves*. 3rd. ed. : The Institution of Engineering and Technology, 2013. p. 335–363.
- BARTELS, J. III - The geomagnetic measures for the time-variations of solar corpuscular radiation, described for use in correlation studies in other geophysical fields. In: INTERNATIONALE, C. I. C. of Scientific Unions Comité Spécial de l'Année G. (Ed.). *Geomagnetism*. : Pergamon, 1957, (Annals of The International Geophysical Year, v. 4). p. 227–236.
- BHUYAN, P. K.; BORAH, R. R. TEC derived from GPS network in India and comparison with the IRI. *Advances in Space Research*, v. 39, no. 5, p. 830–840, 2007.
- BISHOP, G. J.; KLOBUCHAR, J. A.; DOHERTY, P. H. Multipath effects on the determination of absolute ionospheric time delay from GPS signals. *Radio Science*, v. 20, no. 3, p. 388–396, May-June 1985.
- BORRE, K. et al. *A Software-Designed and GPS Galileo Receiver - A Single-Frequency Approach*. : Birkhäuser Boston, 2007.
- BORRE, K.; STRANG, G. *Algorithms for global positioning*. Wellesey MA USA: Cambridge University Press, 2012. 450 p.
- BRAASCH, M. S. Multipath effects. In: PARKINSON, B. W.; JR., J. J. S. (Ed.). *Global Positioning Systems: Theory and Applications*. : American Institute of Aeronautics and Astronautics, 1996, (Progress in Astronautics and Aeronautics, v. 1). ch. 14, p. 547–568.
- BREE, R. J. P. van; TIBERIUS, C. C. J. M. Real-time single-frequency precise point positioning: accuracy assessment. *GPS Solutions*, v. 16, no. 2, p. 259–266, April 2012.
- CESARONI, C. et al. L-band scintillations and calibrated total electron content gradients over Brazil during the last solar maximum. *Journal of Space Weather and Space Climate*, v. 5, no. A36, p. 1–11, December 2015.

- CIRAOLLO, L. et al. Calibration errors on experimental slant total electron content (TEC) determined with GPS. *Journal of Geodesy*, v. 81, p. 111–120, February 2007.
- CLOSAS, P.; FERNÁNDEZ-PRADES, C.; FERNÁNDEZ-RUBIO, J. A. A bayesian approach to multipath mitigation in gnss receivers. *IEEE Journal of Selected Topics in Signal Processing*, v. 3, no. 4, p. 695–706, August 2009.
- CONKER, R. S. et al. Modeling the effects of ionospheric scintillation on GPS/Satellite-Based Augmentation System availability. *Radio Science*, v. 38, no. 1, p. 1–23, January 2003.
- CYBENKO, G. Approximation by superpositions of a sigmoidal function. *Math. Control Signals Systems*, v. 2, p. 303–314, 1989.
- DENARDINI, C. M.; DASSOB, S.; G.-ESPARZAD, J. A. Review on space weather in Latin America. 2. the research networks ready for space weather. *Advances in Space Research*, v. 58, no. 10, p. 1940–1959, March 2016.
- DENARDINI, C. M.; DASSOB, S.; G.-ESPARZAD, J. A. Review on space weather in Latin America. 3. Development of space weather forecasting centers. *Advances in Space Research*, v. 58, no. 10, p. 1960–1967, March 2016.
- DoD. *Interface Specification IS-GPS-200: Navstar GPS Space Segment/Navigation User Interfaces*. 2012.
- DoD, USA. *Global Positioning System Standard Positioning Service Signal Specification*. 2nd. ed. June 1995.
- DYRUD, L. et al. Ionospheric measurement with GPS: receiver techniques and methods. *Radio Science*, v. 43, no. 6, p. 1–11, December 2008.
- EL-HATTAB, A. I. Influence of GPS antenna phase center variation on precise positioning. *NRIAG Journal of Astronomy and Geophysics*, v. 2, no. 2, p. 272–277, December 2013.
- ELMAS, Z. G. et al. Higher order ionospheric effects in GNSS positioning in the european region. *Annales Geophysicae*, v. 29, p. 1383–1399, August 2011.
- ELSOBEIEY, M.; AL-HARBI, S. Performance of real-time Precise Point Positioning using IGS real-time service. *GPS Solutions*, v. 20, no. 3, p. 565–571, July 2015.
- European Space Agency. <http://www.esa.int/Our_Activities/Navigation/Galileo/What_is_Galileo>. December 2017. Galileo web page.
- GLONASS information and Analysis Center. <https://www.glonass-iac.ru/en/GLONASS/>. December 2017. Official Website.
- HABARULEMA, J. B.; MCKINNELL, L.-A.; CILLIERS, P. J. Prediction of global positioning system total electron content using neural networks over South Africa. *Journal of Atmospheric and Solar-Terrestrial Physics*, v. 69, p. 1842–1850, September 2007.
- HABARULEMA, J. B. et al. Application of neural networks to South African GPS TEC modelling. *Advances in Space Research*, v. 43, no. 11, p. 1711–1720, June 2009.
- HADAS, T. et al. Near-real-time regional troposphere models for the GNSS precise point positioning technique. *Measurement Science and Technology*, v. 24, no. 5, p. 1–12, 2013.

- HARTMANN, G. K.; LEITINGER, R. Range errors due to ionospheric and tropospheric effects for signal frequencies above 100 MHz. *Bulletin Geodesique*, v. 58, no. 2, p. 109–136, June 1984.
- HAYKIN, S. *Neural Networks: A Comprehensive Foundation*. 2nd. ed. : Prentice Hall, 1999.
- HAYKIN, S. *Neural Networks and Learning Machines*. 3rd. ed. Bangladesh India: Pearson, 2009. 906 p.
- HOFMANN-WELLENHOF, B.; LICHTENEGGER, H.; WASLE, E. *GNSS - Global Navigation Satellite Systems: GPS, GLONASS, Galileo & more.* : SpringerWienNewYork, 2008. 516 p.
- HOQUE, M. M.; JAKOWSKI, N. Ionospheric propagation effects on GNSS signals and new correction approaches. In: JIN, S. (Ed.). *Global Navigation Satellite Systems: Signal, Theory and Applications.* : InTech, 2012. p. 381–404.
- HOQUE, M. M.; JAKOWSKI, N. An alternative ionospheric correction model for Global Navigation Satellite Systems. *Journal of Geodesy*, v. 89, no. 4, p. 391–406, April 2015.
- JAKOWSKI, N. et al. Total electron content models and their use in ionosphere monitoring. *Radio Science*, v. 46, no. 6, p. 1–11, December 2011.
- JIN, R.; JIN, S.; FENG, G. M DCB: Matlab code for estimating GNSS satellite and receiver differential code biases. *GPS Solutions*, v. 16, p. 541–548, July 2012.
- KAPLAN, E. D.; HEGARTY, C. J. *Understanding GPS : Principles and Applications*. 2nd. ed. : Artech House, 2006.
- KLOBUCHAR, J. A. Ionospheric time-delay algorithm for single-frequency GPS users. *IEEE Transactions on Aerospace and Electronic Systems*, AES-23, no. 3, May 1987.
- KLOBUCHAR, J. A. Ionospheric effects on gps. In: JR., J. J. S. et al. (Ed.). *Global Positioning System: Theory and Applications.* : American Institute of Aeronautics and Astronautics, 1996, (Progress in Astronautics and Aeronautics, v. 1). ch. 12, p. 485–515.
- LANGLEY, R. B. The effect of the ionosphere and troposphere on satellite positioning systems. *Proceedings of Symposium on Refraction of Transatmospheric Signals in Geodesy*, J.C. de Munck, and T.A. Spoelstra (Eds.) Netherlands Geodetic Commission, Publications on Geodesy, New Series No. 36, The Hague, The Netherlands, (abstract). The paper is available on the University of New Brunswick Web server, p. 97, May 1992.
- LANGLEY, R. B. GPS receivers and the observables. In: KLEUSBERG, A.; TEUNISSEN, P. J. G. (Ed.). *GPS for Geodesy.* : Springer-Verlag, 1996. p. 141–173.
- LEANDRO, R. F.; SANTOS, M. C. A neural network approach for regional vertical total electron content modelling. *Studia Geophysica et Geodaetica*, v. 51, no. 2, p. 279–292, April 2007.
- LIU, Z. et al. Influence of higher-order ionospheric delay correction on GPS precise orbit determination and precise positioning. *Geodesy and Geodynamics*, v. 7, no. 5, p. 369–376, May 2016.

MACHADO, W. C. *Utilização de redes neurais artificiais na previsão do VTEC visando a geração de estações de referência virtuais em tempo real*. Thesis (PhD) — Escola Politécnica da Universidade de São Paulo, 2012.

MANUCCI, A. J. et al. A global mapping technique for GPS-derived ionospheric total electron content measurements. *Radio Science*, v. 33, no. 3, p. 565–582, May 1998.

MARINI, J. W. Correction of satellite tracking data for an arbitrary tropospheric profile. *Radio Science*, v. 7, p. 223–231, February 1972.

MARTINS, S. A. M.; NEPOMUCENO, E. G.; BARROSO, M. F. S. Improved structure detection for polynomial NARX models using a multiobjective error reduction ratio. *Journal of Control, Automation and Electrical Systems*, v. 24, no. 6, p. 764–772, September 2013.

MITCHELL, C. The ionosphere. In: BARCLAY, L. (Ed.). *Propagation of Radio Waves*. 3rd. ed. : The Institution of Engineering and Technology, 2013. ch. 11, p. 187–198.

MöLLER, G.; WEBER, R.; BöHM, J. Improved troposphere blind models based on numerical weather data. *Navigation: Journal of the Institute of Navigation*, v. 61, no. 3, p. 203–211, 2014.

MONICO, J. F. G. *Posicionamento pelo GNSS: descrição, fundamentos e aplicações (GNSS positioning: description, fundamentals and applications)*. 2nd. ed. : Editora Unesp, 2008. 476 p.

M.ROTHACHER; G.BEUTLER. The role of GPS in the study of global change. *Physics and Chemistry of the Earth*, v. 23, no. 9-10, p. 1029–1040, 1998.

MYLNIKOVA, A. A. et al. Variability of GPS/GLONASS differential code biases. *Results in Physics*, v. 5, p. 9–10, December 2014.

NIE, Z. et al. An approach to GPS clock prediction for real-time PPP during outages of RTS stream. *GPS Solutions*, v. 22, p. 1–14, January 2018.

Novatel Inc. *An introduction to GNSS - GPS, GLONASS, BeiDou, Galileo and other Global Navigation Satellite Systems*. 2nd. ed. Calgary, Alberta, CA: Novatel Inc., 2015.

OKOH, D. et al. A regional GNSS-VTEC model over Nigeria using neural networks: A novel approach. *Geodesy and Geodynamics*, v. 7, no. 1, p. 19–31, April 2016.

PAPARINI, C. et al. Investigation on the performance of a low-cost single frequency gnss receiver for PPP application. In: *Proceedings of the 19th International Beacon Satellite Symposium*. 2016. p. 3.

ROCHA, G. D. C.; MARQUES, H. A.; GALERA, J. G. M. Acurácia do posicionamento absoluto GPS com correção da ionosfera advinda de mapas ionosféricos globais e regionais. *BCG Boletim de Ciências Geodésicas*, v. 21, no. 3, p. 498–514, 2015.

ROCHA, R. S. M. et al. Avaliação do efeito da cintilação ionosférica e de diferentes intervalos de tempo de coleta de dados no posicionamento por ponto preciso na sua forma on-line. *Revista Brasileira de Geomática*, v. 5, no. 2, p. 251–276, 2017.

SAPUCCI, L. F. *Estimativa do vapor d'água atmosférico e avaliação da modelagem do atraso zenital troposférico utilizando GPS*. Master Thesis (MSc) — Faculdade de Ciências e Tecnologia UNESP, 2001.

- SEEBER, G. *Satellite Geodesy*. 2nd. ed. : Walter de Gruyter GmbH & Co., 2003.
- SEEMALA, G. K. *GPS-TEC analysis application*. 2017.
- SPIPKER, J. J. Tropospheric effects on GPS. *Global Positioning System: Theory and Applications*, v. 1, p. 517–546, 1996.
- TAKAHASHI, H. et al. Diagnostics of equatorial and low latitude ionosphere by TEC mapping over Brazil. *Advances in Space Research*, v. 54, no. 3, p. 385–394, August 2014.
- TAPPING, K. F. The 10.7 cm solar radio flux ($f_{10.7}$). *Space Weather*, v. 11, p. 394–406, July 2013.
- TEUNISSEN, P. J. G.; KLEUSBERG, A. Propagation of the GPS signals. In: KLEUSBERG, A.; TEUNISSEN, P. J. G. (Ed.). *GPS for Geodesy*. : Springer-Verlag, 1996. p. 174–217.
- The European Commission. *Open Service signal-in-space interface control document*. December 2016.
- TIWARI, R. et al. WMod assisted PLL GPS software receiver for mitigating scintillation affect in high latitude region. In: INTERNATIONAL UNION OF RADIO SCIENCE. *2011 XXXth URSI General Assembly and Scientific Symposium*. Istanbul, 2011. p. 1–4.
- US Department of Defense. <https://www.gps.gov/systems/gps/space/>. 2017. Official GPS web-site.
- WEBSTER, I. R. *A Regional Model for the Prediction of Ionospheric Delay for Single Frequency Users of the Global Positioning System*. Fredericton, New Brunswick, Canada, 1993.
- WEINBACH, U.; SCHÖN, S. GNSS receiver clock modeling when using high-precision oscillators and its impact on PPP. *Advances in Space Research*, v. 47, no. 2, p. 229–238, January 2011.
- WILLISCROFT, L. A.; POOLE, A. W. V. Neural networks, foF2, sunspot number and magnetic activity. *Geophysical Research Letters*, v. 23, no. 24, p. 3659–3662, December 1996.
- YAO, Y. et al. ITG: A new Global GNSS Tropospheric Correction Model. *Scientific Reports*, v. 5, no. 10273, July 2015.
- YIBIN, Y. et al. A global empirical model for estimating zenith tropospheric delay. *Science China Earth Sciences*, v. 59, no. 1, p. 118–128, January 2016.
- YU, X. et al. Ionospheric correction based on ingestion of global ionospheric maps into the Nequick 2 model. *The Scientific World Journal*, v. 2015, no. 376702, 2015.
- ZHAO, J.; HAN, Y. B. Historical dataset reconstruction and a prediction method of solar 10.7 cm radio flux. *Chinese Journal of Astronomy and Astrophysics*, v. 8, no. 4, p. 472–476, January 2008.

Appendices

Appendix A

List of publications

A.1 Directly related publications

FERREIRA, A. A.; BORGES, R. A.; PAPANINI, C.; CIRAOLO, L.; RADICELLA, S. M.; *Short-term estimation of GNSS TEC using a neural network model in Brazil*. Advances in Space Research, v. 60, p. 1765-1776, 2017.

FERREIRA, A. A.; BORGES, R. A.; PAPANINI, C.; RADICELLA, S.; *TEC modelling via neural network using observations from the first GLONASS R&D data network in Brazil and the RBMC*. Proceedings of the 20th World Congress of the International Federation of Automatic Control (IFAC), Toulouse, 2017. p. 13371-13376.

FERREIRA, A. A.; BORGES, R. A.; Análise de diferentes estruturas de redes neurais artificiais Multilayer Perceptron na modelagem do conteúdo total de elétrons.(Analysis of different structures of Multilayer Perceptron artificial neural networks in the modelling of the Total Electron Content.) Presented at the XIII Brazilian Symposium on Intelligent Automation, 2017 (in Portuguese).

A.2 Indirectly related publications

PAUL, A.; KASCHEYEV, A.; RODRIGUEZ-BOUZA, M.; PATHAK, K.; **FERREIRA, A.A.**; SHETTI, D.; YAO, J.N. Latitudinal features of Total Electron Content over the African and European longitude sector following the St. Patrick's day storm of 2015. Advances in Space Research, 2017 (*in press*).

Appendix B

Linearization of the observable equation

By using the ephemeris sent in the navigation message it is possible to obtain the satellite positioning coordinates (X^{sv}, Y^{sv}, Z^{sv}) and the satellite clock bias dt^{sv} . The terms corresponding to ionospheric and tropospheric delays $(I_r^{sv} \text{ e } T_r^{sv})$ can be obtained from *a priori* models (BORRE et al., 2007). Some strategies to deal with multipath effect (dm_r^{sv}) are presented in Section 2.3.2. Thus, the unknowns of Equation (2.5) are the receiver position coordinates (X_r, Y_r, Z_r) and the receiver clock bias.

Since Equation (2.5) is non-linear with respect to the coordinates of the receivers and satellites, it has to be linearized before applying the least-squares method (BORRE et al., 2007). Considering the non-linear term of the Equation (2.5) it follows, for an epoch t:

$$\rho_r^{sv}(t) = \sqrt{(X^{sv}(t) - X_r)^2 + (Y^{sv}(t) - Y_r)^2 + (Z^{sv}(t) - Z_r)^2}. \quad (\text{B.1})$$

By assuming the approximated values $(X_{r,0}, Y_{r,0}, Z_{r,0})$ for the coordinates of the receiver (station) r , an approximated distance can be obtained by:

$$\rho_r^{sv}(t) = \sqrt{(X^{sv}(t) - X_{r,0})^2 + (Y^{sv}(t) - Y_{r,0})^2 + (Z^{sv}(t) - Z_{r,0})^2}, \quad (\text{B.2})$$

and the coordinates of the station can be represented by:

$$\begin{aligned} X_r &= X_{r,0} + \Delta X_r \\ Y_r &= Y_{r,0} + \Delta Y_r \quad , \\ Z_r &= Z_{r,0} + \Delta Z_r \end{aligned} \quad (\text{B.3})$$

where ΔX_r , ΔY_r and ΔZ_r correspond to the corrections to the approximated values (MONICO, 2008). This initial approximated coordinate is often chosen as the center of the Earth in the ECEF coordinate system (0,0,0) (BORRE et al., 2007). By applying (B.3) in (B.2) and

expanding the resulting expression as a first-order Taylor series, one can obtain:

$$\rho_r^{sv}(t) = \rho_{r,0}^{sv}(t) + \frac{\partial \rho_{r,0}^{sv}(t)}{\partial X_{r,0}} \Delta X_r + \frac{\partial \rho_{r,0}^{sv}(t)}{\partial Y_{r,0}} \Delta Y_r + \frac{\partial \rho_{r,0}^{sv}(t)}{\partial Z_{r,0}} \Delta Z_r. \quad (\text{B.4})$$

The partial derivatives presented in Equation (B.4) are (MONICO, 2008; BORRE et al., 2007):

$$\begin{aligned} \frac{\partial \rho_{r,0}^{sv}(t)}{\partial X_{r,0}} &= -\frac{X^{sv} - X_{r,0}}{\rho_{r,0}^{sv}(t)}, \\ \frac{\partial \rho_{r,0}^{sv}(t)}{\partial Y_{r,0}} &= -\frac{Y^{sv} - Y_{r,0}}{\rho_{r,0}^{sv}(t)}, \\ \frac{\partial \rho_{r,0}^{sv}(t)}{\partial Z_{r,0}} &= -\frac{Z^{sv} - Z_{r,0}}{\rho_{r,0}^{sv}(t)}. \end{aligned} \quad (\text{B.5})$$

Now the Equation (B.4) is linear with respect to the unknowns ΔX_r , ΔY_r , ΔZ_r and the pseudorange equation can be written as:

$$\begin{aligned} P_r^{sv} = \rho_{r,0}^{sv} - \frac{X^{sv} - X_{r,0}}{\rho_{r,0}^{sv}} \Delta X_r - \frac{Y^{sv} - Y_{r,0}}{\rho_{r,0}^{sv}} \Delta Y_r - \frac{Z^{sv} - Z_{r,0}}{\rho_{r,0}^{sv}} \Delta Z_r \\ + c(\delta t_r - \delta t^{sv}) + T_r^{sv} + I_r^{sv} + dm_r^{sv} + \varepsilon_P. \end{aligned} \quad (\text{B.6})$$

By using the linearized pseudorange equation, it is possible to apply the least-squares method to perform the receiver positioning calculation.

Appendix C

Additional tropospheric effect information

C.1 Tropospheric attenuation as a function of the elevation angle

By modelling the troposphere as a simple and uniform spherical shell of height h_m above the Earth, the length of the signal path L through the troposphere varies with the elevation angle (ϵ) and is given by (SPILKER, 1996)

$$L = \frac{2h_m(1 + a/2)}{\sin(\epsilon) + \sqrt{\sin^2(\epsilon) + 2a + a^2}} \quad (\text{C.1})$$

where $h_m = 6$ km, R_e corresponds to the Earth radius ($R_e \approx 6378$ km) and $a = h_m/R_e$. Figure C.1 illustrates the length L of signal path in which $\cos(\epsilon) = (1 + a)\cos(\phi)$.

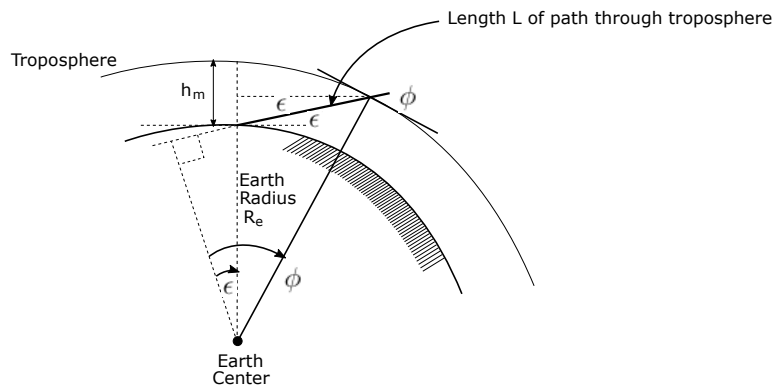


Figure C.1: Path length L through an uniform shell troposphere at elevation angle ϵ (SPILKER, 1996) (*adapted*).

The oxygen attenuation $A(\epsilon)$ has the following approximate value (SPILKER, 1996)

$$A(\epsilon) \approx \frac{2A(90^\circ)(1 + a/2)}{\sin(\epsilon) + \sqrt{(\sin^2(\epsilon) + 2a + a^2)}} \approx \begin{cases} \frac{2A(90^\circ)}{\sin(\epsilon)+0.043} \text{ dB}, & \text{for small } \epsilon \text{ but greater than } 3^\circ \\ \frac{A(90^\circ)}{\sin(\epsilon)} \text{ dB}, & \text{for } \epsilon \text{ but greater than } 10^\circ \end{cases} \quad (\text{C.2})$$

where $a = h_m/R_e$, $h_m = 6\text{km}$ and R_e corresponds to the Earth radius ($R_e \approx 6378\text{km}$).

To illustrate the attenuation effect and its relation to the elevation angle the attenuation presented in Equation (C.2) is plotted in Figure C.2 for elevation angles greater than 3° .

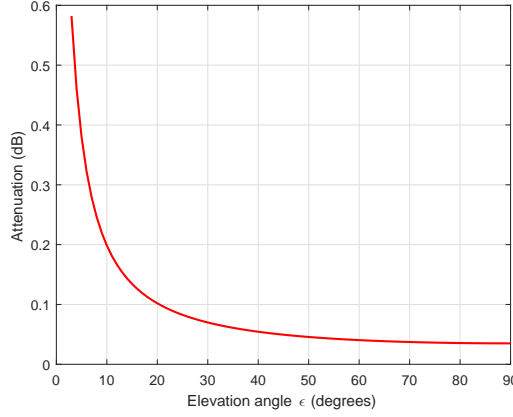


Figure C.2: (Tropospheric attenuation versus elevation angle (SPILKER, 1996) (*adapted*)).

Near the horizon, e.g. below 3° , the uniform spherical model of Figure C.1 is no longer accurate, and neither Equation (C.2) and Figure C.2 should be used (SPILKER, 1996). In addition to attenuation potential effects such as multipath and refraction can be increased at low elevation angles (SPILKER, 1996).

C.2 Other tropospheric effects

Rainfall attenuation

The effect of rainfall attenuation on GNSS signals is very small. Considering a dense rainfall (100 mm/h) and a signal with frequency equals to 2 GHz, the attenuation due to this event is less than 0.01 dB/km. For signals with frequency below 2 GHz, the effect is even smaller (SPILKER, 1996).

Tropospheric scintillation

The tropospheric scintillation is caused by irregularities and turbulence in the atmospheric refractive index mainly in the first few kilometers above the ground (SPILKER, 1996). A GNSS signal propagating through the troposphere is affected by a combination of

random absorption and scattering. These effects cause variations in the amplitude and phase in the received waveform (MONICO, 2008; SPILKER, 1996). Tropospheric scintillation effect varies with time and depends on signal frequency, elevation angle, and weather conditions, particularly dense clouds. For elevation angles higher than 10° the most significant effect is the scattering caused by the turbulence. In general, for GNSS signals these effects are small, except for lower elevation angles (SPILKER, 1996; MONICO, 2008).

C.3 Empirical models and mapping functions

The empirical models for tropospheric delay are functions of meteorological parameters, such as temperature, pressure and humidity. These parameters allow determining the Zenith Troposphere Delay (ZTD), which is mapped into Slant Tropospheric Delay (STD) by using the mapping functions (HADAS et al., 2013).

As presented in Section 2.3.3.2 it is necessary to obtain the refractivity index N_T along the path between the receiver and the satellite to calculate the tropospheric delay from Equation (2.10). The empiric general expression for the refractivity of a non-ideal gas, including water vapor, is given by (MONICO, 2008):

$$N_T = k_1 \frac{P_h}{T} Z_H^{-1} + k_2 \frac{e}{T} Z_W^{-1} + k_3 \frac{e}{T^2} Z_W^{-1} \quad (\text{C.3})$$

where:

- P_h and e are the partial pressure of the dry air and water vapor (expressed in millibars);
- T is the temperature expressed in Kelvin;
- Z_W and Z_H corresponds to the compressibility factors for the wet and hydrostatic components, respectively;
- the constants k_1, k_2 e k_3 are obtained experimentally with $k_1 = 77.604 \pm 0.0124$, $k_2 = 64.79 \pm 10$ and $k_3 = 377600 \pm 3000$.

By using the gases law, one can obtain the following expressions for the hydrostatic and wet refractivities,

$$N_H = \frac{k_1 R}{M_H} \rho = 22.276 \rho, \quad (\text{C.4})$$

$$N_W = \left[k_2 - k_1 \frac{M_W}{M_H} \right] \frac{e}{T} Z_W^{-1} + k_3 \frac{e}{T^2} Z_W^{-1}. \quad (\text{C.5})$$

The hydrostatic refractivity presents a low uncertainty ($\pm 0.014\%$) and depends only on the total density of the atmosphere, therefore it can be determined with good precision. The wet refractivity, on the other hand, presents considerable uncertainty in its coefficients. This fact difficults the N_W determination (MONICO, 2008).

Since the refractivity along the way between the receiver antenna and the satellite is very complicated to obtain, several models have been developed to perform *a priori* corrections of the tropospheric delay. In these models, additional parameters can be included in the form of adjustment parameters. Some common models employed on tropospheric delay correction are the *Hopfield model* and Standard Model.

A brief description of these models are as follows:

- a) *Hopfield model*: Among the most knowns models for tropospheric delay, this model was developed in the decade of 1960 and is given by the following set of equations:

$$\begin{aligned}
T_r^s &= T_{ZH} \times mh(\epsilon) + T_{ZW} \times mw(\epsilon) \\
T_{ZH} &= 155.2 \times 10^{-7} \frac{P}{T} H_d \\
T_{ZW} &= 155.2 \times 10^{-7} \frac{4819e}{T^2} H_w \\
mh(\epsilon) &= \frac{1}{\sqrt{\epsilon^2 + 6.25}} \\
mw(\epsilon) &= \frac{1}{\sqrt{\epsilon^2 + 2.25}} \\
H_d &= 40136 + 148.72(T - 273.16) \\
H_w &= 11000 \text{ m}
\end{aligned} \tag{C.6}$$

where, the terms T_{ZH} , T_{ZW} , $mh(\epsilon)$ e $mw(\epsilon)$ are the zenith tropospheric components and the mapping functions described previously in Section (2.3.3.2), and ϵ is the elevation angle. Other mapping functions can also be employed.

- b) *Saastamoinen model*: In this model, it is assumed a linear decreasing in the temperature until tropopause (about 12 km height); for higher altitudes the stratosphere is characterized as an isothermal model by a constant value. Other assumptions were made in order to develop this model, such as: *i*) atmosphere in hydrostatic equilibrium, *ii*) the water vapor is concentrated in the troposphere and behaves as an ideal gas. The standard model proposed by Saastamoinen 1972, with some adjustments, is given by (MONICO, 2008; SPILKER, 1996):

$$T_r^s = 0.002277 \frac{1}{1 + D} \frac{1}{\cos z} \left[P + \left(\frac{1255}{T} + 0.05 \right) e - B \tan^2 z \right] + \partial_R, \tag{C.7}$$

where P and e represent the superficial pressure (expressed in mBar) of dry air and water vapor, respectively; T is the temperature in Kelvin; B and ∂_R are correction factors and $z = 90^\circ - \epsilon$. The values of B and ∂_R for various users heights can be found in Spilker 1996. The value D is obtained by

$$D = 0.0026 \cos(2\phi) + 0.00028H, \tag{C.8}$$

where ϕ corresponds to the latitude and H is the orthometric height in km.

The values of atmospheric pressure and water vapor pressure (P , e) are expressed in mBar. The temperature value T is expressed in Kelvin and usually is obtained from a standard atmosphere. Hence, the following values can be employed

$$\begin{aligned} P &= P_r(1 - 0.0000226(H - H_r))^{5.225}, \\ T &= T_r - 0.0065(H - H_r), \\ H_u &= H_{ur}e^{-0.0006396(H-H_r)}, \end{aligned} \quad (\text{C.9})$$

where P_r , T_r and H_{ur} are the reference values for pressure, temperature and humidity at the reference height H_r (MONICO, 2008).

In addition to the models presented previously, some *blind models*, which relies on climatological data have been proposed, such as (YIBIN et al., 2016; MÖLLER; WEBER; BÖHM, 2014; YAO et al., 2015):

- RTCA-MOPS;
- ESA model;
- GPT2;
- GPT2w;
- UNB3, UNB3m, UNB4;
- ITG.

As presented before, it is necessary to use mapping functions in order to convert the zenith tropospheric delay to the slant tropospheric delay. In some models, such as *Saastamoinen*, the mapping functions are implicit in the model. One simple mapping function to be used is

$$mh(\epsilon) = \frac{1}{\sin(\epsilon)}. \quad (\text{C.10})$$

This mapping function can be a good approximation for elevation angles near to the zenith ($\epsilon \approx 90^\circ$). For ϵ below this value, it is necessary to use more sophisticated approximations.

Among the available mapping functions, one can use the Marini mapping function and its derivatives, such as Chao, Niell, Vienna, Global Mapping Function, and others (MONICO, 2008).

The mapping function proposed by Marini is given by (MARINI, 1972):

$$m(\epsilon) = \frac{1}{\sin(\epsilon) + \frac{a}{\sin(\epsilon) + \frac{b}{\sin(\epsilon) + \frac{c}{\sin(\epsilon)} + \dots}}} \quad (\text{C.11})$$

The Chao mapping function defines the hydrostatic and wet components as follow:

$$\begin{aligned} m_H(\epsilon) &= \frac{1}{\sin(\epsilon) + \frac{0.00143}{\tan(\epsilon)+0.0445}}, \\ m_W(\epsilon) &= \frac{1}{\sin(\epsilon) + \frac{0.00035}{\tan(\epsilon)+0.017}}. \end{aligned} \tag{C.12}$$

Appendix D

Additional ionospheric effect information

Ionospheric scintillation

Plasma density irregularities density along the path of the signal through the ionosphere causes the ionosphere scintillation phenomena. This phenomenon consists of rapid changes in the phase and/or amplitude of the received radio signal. It can reduce the accuracy of GNSS receiver pseudorange and carrier phase measurements and also can result in a loss of lock on a satellite. In a worst-case scenario, it can lead to a loss in the positioning service, depending on the number of satellites involved in the loss of lock (CONKER et al., 2003).

The scintillation effects are more severe in the equatorial region of the Earth, followed by high-latitude (mainly on the poles) and middle latitudes (rare phenomenon, mainly occurring in the presence of ionospheric storms). In equatorial regions, this effect is related to the Equatorial Anomaly. However, the formation of large-scale ionospheric bubbles can increase the occurrence of scintillation in this region. In years near the peak of the solar cycle, this effect can be observed on a daily basis. The periods of equinox tend to show the strongest scintillations. In equatorial regions, the ionospheric scintillation tends to occur in the F-Layer region at the time when the sun sets. (ROCHA et al., 2017; ARBESSER-RASTBURG; ROGERS, 2013).

Several modelling activities have been made to describe the physical mechanism that origins scintillation as well statistical distribution of ionospheric density irregularities, such as WBMOD and GISM models, respectively. The measure of scintillation can be done by the phase scintillation index σ_Φ (given in radians) and the amplitude scintillation index S_4 . Table D.1 presents the three categories of scintillation according to the σ_Φ and S_4 values (TIWARI et al., 2011).

Since a large part of Brazil is situated in the geomagnetic equator, it suffers significant scintillation effect. The highest values of scintillation indexes occur between September to March, approximately. The minimum values generally occur from April to August (ROCHA et al., 2017).

Table D.1: Classification of the scintillation levels.

Classification	S_4	σ_Φ
Strong Scintillation	$S_4 > 1.0$	$\sigma_\Phi > 0.8$
Moderate Scintillation	$0.5 \leq S_4 \leq 1.0$	$0.4 \leq \sigma_\Phi \leq 0.8$
Weak Scintillation	$0 \leq S_4 < 0.5$	$0 \leq \sigma_\Phi < 0.4$

Appendix E

Other sources of Errors

Satellite related errors

In addition to the satellite clock error described in Section 2.3.1.1, this section presents other GNSS errors related to the satellites. These errors are the orbit errors, differential code biases, and antenna phase center.

- a) *Orbit errors* - In order to obtain position and velocity of the GNSS satellites in a terrestrial frame, the following sets of data are available: almanac data, broadcast ephemeris, precise ephemeris and predicted by the International GNSS Service (IGS). These data differ in accuracy and availability to the user (real-time or with some delay). Essentially, the ephemeris contain records with general information, records with orbital information, and records with information on the satellite clock. For the GPS the orbital information is provided in the form of Keplerian parameters together with their temporal variations, whereas for the GLONASS the orbital information is provided in form of position and velocity vectors at equidistant epochs (HOFMANN-WELLENHOF; LICHTHENEGGER; WASLE, 2008).

The satellite coordinates calculated from the ephemeris are in general considered fixed during the process of adjustments of the satellite data. Therefore, the presence of errors on satellite position (coordinates) will be directly transmitted to the user positioning. For precise point positioning (PPP), the errors are almost directly propagated to the user, however, in the relative positioning, the orbital errors are practically eliminated. As the distance of the baseline becomes longer, the remaining errors can influence and degrade the accuracy of the base-line components (MONICO, 2008). A more detailed description about the PPP and relative positioning can be found in Appendix F.

The purpose of the almanac data is to provide the user with adequate data to facilitate receiver satellite acquisition and for planning tasks such as the computation of visibility charts. These data are regularly updated and sent as part of the satellite message. It contains information for orbit and satellite clock correction (HOFMANN-

WELLENHOF; LICHTENEGGER; WASLE, 2008).

The broadcast ephemeris are based on observations occurred at the monitoring stations of the respective control segment. The most recent of these data are used to compute reference orbit for the satellites. In essence, the ephemerides contain records with general information, records with orbital information, and records with information on the satellite clock. The real-time ephemeris transmitted with the observations can reach accuracy from 1 to 3 m (MONICO, 2008; HOFMANN-WELLENHOF; LICHTENEGGER; WASLE, 2008). The precise ephemeris are result of post-processing and can reach accuracy from 2 to 5 cm. Table E.1 presents the uncertainty related to the ephemeris according to its type.

Table E.1: Uncertainty of types of ephemeris (HOFMANN-WELLENHOF; LICHTENEGGER; WASLE, 2008) (*adapted*).

Ephemeris type	Uncertainty	Remark
Almanac	Some kilometres	Depending on the age of data
Broadcast Ephemeris	≈ 1 m	Or even better
Precise Ephemeris	0.025 to 0.20 m	Depending on the latency

- b) *Differential Code Biases (DCBs)* - This error, also called *Interfrequency Biases (IFB)* is associated with the different frequency dependent processing times of L1 and L2 signals in RF paths, both for satellites and receivers (MYLNIKOVA et al., 2014; MONICO, 2008). DCB values vary for each satellite and ground receivers. The DCB values vary between different GNSS satellites and ground receivers. Most DCB estimates are based on the assumption that the DCB values of GPS satellites or receivers are constant over 1 day or 1 month, while in fact, they are often changing in hours or 1 day (JIN; JIN; FENG, 2012). The DCB information is routinely produced by the IGS analysis centers, such as the Center for Orbit Determination in Europe (CODE), the Jet Propulsion Laboratory (JPL), the European Space Agency (ESA), and the Polytechnic University of Catalonia (UPC). (JIN; JIN; FENG, 2012; MONICO, 2008).

The Differential Code Biases has to be taken into account for precise Total Electron Content (TEC) estimations. If DCBs effects are ignored, several meters of error can occur and it can lead to non-physical negative values of TEC (MYLNIKOVA et al., 2014; JIN; JIN; FENG, 2012).

- c) *Relativity effects* - The satellite clock changes due to the effects of general and special relativity. The clocks of the receivers in the ground stations and the clocks aboard the satellites are under influence of different gravitational fields and are moving at distinct velocities. These facts cause an apparent frequency changing in the on-board clocks with respect to the ground clocks. On the GPS, this problem is solved by a reduction on the clock satellite of 4.55×10^{-3} Hz before launching (MONICO, 2008).

The correction term with respect to the relative effect (Δt_R) is shown in Equation (2.9).

In relative positioning (presented in Appendix F), the relativity effects are totally negligible (MONICO, 2008).

- c) *Antenna Phase Center (APC)* - The GNSS antenna is the connecting element between the satellite and the GNSS receiver. The satellite Antenna Phase Center (APC) is the reference point for the broadcast ephemeris. This point differs from the center of mass of the satellite, and it changes for different values of elevation and azimuth angles. Therefore, for high precision applications, the phase center offset vector has to be taken into account (EL-HATTAB, 2013; MONICO, 2008; DoD, USA, 1995).

The APC offset vectors, in general, are given with respect to a fixed coordinate frame, with the origin on the satellite center of mass, and shall be converted to a fixed reference frame on Earth, such as ITRF2005. These values are calculated by some IGS analysis centers (e.g. CODE and GFZ) and are provided by IGS (MONICO, 2008).

Receiver related errors

- a) *Interchannel errors* - Nowadays, the majority of geodesic receivers present multiple channels and in this case, a systematic interchannel error can occur, since the signal of each satellite will travel through different electronic paths. In order to correct this effect, the receivers have a device which calibrates and compensates these interchannel biases (SEEBER, 2003; MONICO, 2008). A channel is taken as the reference, each channel tracks, simultaneously, a particular satellite and determines the errors with respect to the reference channel. This effect will be correct in all subsequent observations.

Appendix F

Positioning techniques

Positioning corresponds to the determination of positions of objects with respect to a specific reference. It can be divided into absolute positioning, when the calculated coordinates are directly associated with the centre of the Earth, and relative positioning, when the coordinates are obtained with respect to an object with known coordinates. Also, the positioning can be classified in static and cinematic, depending on if the object is at rest or under movement (MONICO, 2008).

In the absolute positioning (also named *Point Positioning*), when the transmitted ephemeris are employed, the position of the point is obtained in the reference frame of the system (WGS 84 for GPS, PZ90 for GLONASS). However, when precise ephemeris and clock corrections are employed, along with carrier phase data, the reference frame adopted is related to the precise ephemeris. In addition to the classifications presented previously, the positioning can be classified as real-time and post-processed. In the first, the estimative occurs almost at the same time the observations are collected. In the post-processed, the estimates are obtained after the data collect (MONICO, 2008).

In the following sections the Point Positioning (PP), Precise Point Positioning (PPP) and Relative Positioning (RP) techniques will be presented.

Point Positioning

Point positioning refers to the determination of the position of a station based on pseudo-range observation, using orbit and satellite parameters broadcast in the navigation message. This positioning technique requires only a receiver and is a widely employed method for low precision navigation. If available, it is possible to include in the processing, besides the pseudorange, the carrier phase measurements. However, this procedure is not very common in this technique, since it does not provide an improvement in the solution for one epoch (MONICO, 2008). Therefore, this positioning technique does not satisfy the precision requirements for geodetic positioning.

The major errors that affect this positioning method are related to the quality of the observable (*pseudorange*) and the accuracy of the parameters sent in the navigation message, such as ephemeris and clock. In addition, this method suffers influence of tropospheric and ionospheric errors, multipath, and others. When real-time positioning is not necessary, it is possible to use precise ephemeris and satellite clock corrections, leading to a more precise solution. This technique is called *Precise Point Positioning - PPP* and will be briefly presented in the following section (MONICO, 2008).

Precise Point Positioning

Precise Point Positioning (PPP) is a positioning method similar to the Point Positioning. However, the use of precise satellite orbits and satellite clock corrections provides a better quality to this method (ROCHA et al., 2017; MONICO, 2008). Besides satellite orbit and clock corrections, other errors that affect the observations, such as ionospheric and tropospheric effects, have to be corrected in order to allow a more accurate PPP solution (ROCHA et al., 2017). The types of ephemeris provided by the IGS are presented in Appendix E.

This method presents a great potential to be used in applications that require high-accuracy, such as geodynamics, and presents great advantages when compared with the processing of GNSS networks, in which high-computational effort is required (MONICO, 2008). Given the spread of internet use, some PPP *on-line* services are available to the users (for free in some cases). These *on-line* services offer the positioning estimates, in the post-processing mode, from GNSS observation sent by the users. Several *on-line* services are available, such as the Automatic Precise Positioning Service (APPS) provided by the Jet Propulsion Laboratory (JPL), the MagicGNSS developed by the GMV Aerospace and Defense S.A., the Canadian Spatial Reference System (CSRS-PPP) provided by the Natural Resource Canada - NRCAN. In Brazil, national initiatives such as the RT-PPP developed at the School of Sciences and Technology of São Paulo State University (FCT/UNESP) and the IBGE-PPP provided by the Brazilian Institute of Geography and Statistics (IBGE) are also available to the users who want to perform the on-line PPP solutions (ROCHA et al., 2017).

In addition to the post-processed PPP, the real-time PPP solution has been investigated during the last years, achieving good accuracy results. In 2009, van Bree et al. 2009 compared the results obtained in the real-time Single-Frequency PPP solution using IGS Ultra-Rapid products (satellite clock error) and the REal-Time CLock Estimation (RETICLE) developed by at German Space Operations Center of the German Aerospace Center (GSOC/DLR) with the IGS Final clocks. The results obtained with the RETICLE satellite clock presented an accuracy comparable with the IGS Final Products. In the work (BREE; TIBERIUS, 2012), the authors assessed the SF-PPP solution using satellite clocks and orbits provided by RETICLE, and used GIMs and predicted DCB from the Center for Orbit Determination (CODE). The position accuracy achieved with a kinematic approach reached

standard deviations of about 0.15 and 0.30 m for the horizontal and vertical coordinates, respectively, and with 95% error values of about 0.30 m and about 0.65 m, where the 95th percentile means that 95% the error is below this amount.

The international GNSS service (IGS) has been providing an open-access real-time service (RTS) since 2013. The advent of IGS RTS products allows users to carry out real-time precise point positioning (RT-PPP) anywhere in the world (NIE et al., 2018). In the study carried out by (ELSOBEIEY; AL-HARBI, 2015), the performance of RT-PPP using the IGS RTS is assessed. The accuracy obtained using the RTS products presented better accuracy when compared to the IGS ultra-rapid products (predicted part). In addition to these studies on Real-Time PPP, a recent study carried out by (NIE et al., 2018) proposes an alternative approach for GPS clock prediction in case of an outage on RTS. These investigations in Real-Time PPP can improve applications such as high-precision agriculture, navigation applications in the automotive and possibly even automated aircraft landing (BREE; TIBERIUS, 2012).

Relative Positioning

The relative positioning involves the simultaneous observation of a number m of satellites and a minimum of two GNSS-receivers (TEUNISSEN; KLEUSBERG, 1996). However, with the advent of the Active Control Systems (ACS) it is possible to perform the relative positioning by using only one GNSS receiver. In an ACS, the receivers perform continuous tracking of the visible satellites and the users can access the data via communication system (MONICO, 2008).

The advantage of the relative positioning is related to the fact that the parameters of interest are much less sensitive to interfering uncertainties such as ephemeris, clocks and atmospheric effects (TEUNISSEN; KLEUSBERG, 1996). By combining the observables involved in the process of relative positioning it is possible to mitigate some systematic errors that affect the GNSS observables errors, such as atmospheric effects and satellite clock errors.