

**Modeling and prediction of ionospheric characteristics using
nonlinear autoregression and neural network**

by

Hendy Santosa

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF ENGINEERING

at

The University of Electro-Communications

非線形同定手法を用いた電離層特性のモデリングおよび予測に関する研究

ヘンデイサントーサ

論文の和文要旨

地球を取り巻く電離層は、その中を通過、反射する電磁波の伝搬特性に多大な影響を及ぼすことが知られている。電離層の継続的な特性の観測には、電磁波による遠隔探査（リモートセンシング）が広く用いられている。例えば、電離層中で最大の電子密度をもつ F 層の観測には、電磁波の鉛直打ち上げ（イオノゾンデ）が用いられ、電子密度の高度分布が得られる。一方、MLT（Mesosphere-Lower Thermosphere）領域である電離層下端の D 層は、その電子密度の小ささゆえ、VLF 帯送信電波の受信振幅、位相の変化がほぼ唯一の観測方法である。電離層は、その上層及び下層起源の様々な要因（外因）により、時間空間的に複雑に変動する。例えば、上層からの外因として太陽活動の影響が挙げられる一方で、下層からの外因として大気の影響が挙げられる。しかしながら、外因別の電離層の変動への定量的評価は、現在までにほとんど行われていない。そこで本研究では、非線形システム同定手法を用いて、D 層および F2 層における、電離層の時間変化の予測モデルを構築し、外因別の貢献度を導出した。特に、D 層の電離層状態を表す VLF 送信電波振幅の時間変動予測モデルの構築は初の試みである。さらに、D 層、F 層ともに、外因別の貢献度の導出も初めてである。今後これらの研究成果は、電離層のダイナミクスの理解、太陽活動や大気パラメータの監視、通信障害の予測、さらには地震に関連する異常値検出等に貢献する可能性がある。D 層に関しては、電通大 VLF 帯送信電波観測ネット

ワークにより受信された様々な緯度を通過する国内外からの送信電波の電界振幅の長期時系列観測データを解析した。まず、電界振幅データに非線形システム同定手法の一つである NARXNN (Nonlinear AutoRegressive with eXogenous input Neural Network) を適用して、プロセスモデルを生成し、電界振幅の時間変化に対する支配的な外因を同定した。次に、これらの支配的な外因を用いて、予測を行い構築されたモデルの評価を実施した。その結果、予測値と観測値の間に非常に高い相関係数が得られた。また、緯度の異なる伝搬経路において支配的な外因に違いが見られ、その理由に対する物理的考察を行った。さらに、F 層に関しては、中緯度帯である日本国内で観測された、イオノゾンドのデータを用いて、D 層と同様の解析を実施し、予測モデルを構築するとともに、D 層と F 層間の特性の相違を調査した。

Modeling and prediction of ionospheric characteristics using nonlinear autoregression and neural network

Hendy Santosa

Abstract

The terrestrial ionosphere from D-region (60 km) to F-region (500 km) plays an important role in radio wave propagation between the Earth and ionosphere. During the last half-century, a considerable experimental, theoretical, and modeling efforts have been made to understand the physical process occurred in the ionosphere at different altitudes. Radio sensing techniques is widely used to continuously monitor the ionospheric conditions. For example, the ionospheric property in the F2 layer is obtained by a vertical sounding so-called Ionosonde. Properties of the D layer (the lower end of the ionosphere) is effectively obtained by receiving VLF/LF transmitter signals. Although, the ionospheric condition varies both in time and space due to various external forcings from the atmosphere and space weather parameters, quantitative information of contributions influencing the ionosphere from every external forcing have not understood well. In this thesis nonlinear autoregressive with exogenous input and neural network is applied first time to identify the ionospheric characteristics based on the VLF radio wave propagation and ionosonde. One step ahead prediction of the daily nighttime means of VLF electric amplitude in three different latitude paths and two receiving stations by using NARXNN has been carried out. The relative contribution to the ionospheric conditions (VLF electric amplitude variability) from every external forcing has been revealed. Moreover, the proposed model extends for multi-step ahead prediction to evaluate the performance of prediction accuracy for five and ten days ahead. The temporal dependence of F2-region critical frequency (f_oF_2) has been predicted by using the same approach as used for the VLF signals. Physical interpretation of relative contribution to the ionospheric conditions from major external forcing parameters have been made. The results of this thesis can be used to detect anomalies in relation with severe weather, major seismic activity, and space weather to mitigate damages and human victims. Furthermore, we investigate the coupling from external sources between the D- and F-region in the middle-latitude path.

Acknowledgements

I am very grateful to my supervisor, Prof. Yasuhide Hobara, for his indispensable advice at every step helping me to make my work conceptually sound and also to draw my attention to even the smallest mistakes to improve my academic writing. With his great enthusiasm and efforts to explain things clearly and simply, he has guided me to the point where I have been able to complete the work of my thesis.

I will remain in debt to both my co-supervisors Prof. Yanagisawa and Prof. Yoshiaki Ando, Prof. Balikhin and my laboratory assistant professor Dr. Takuo Tsuda for their priceless advice that I have progressed towards becoming a professional researcher. I would like to thank my friends, Dr. Sujay Pal, Dr. Tamal Basak, Dr. Satya Vemuri Srinivas, Mr. Tomoki Kawano, Mr. Yuma Matsui, Mr. Katsunori Suzuki and Mrs. Emiko Takahashi, for assisting me in preparing a gold standard for the evaluation of my results. I would like to thank all my colleagues for sharing their valuable experience to guide me through writing the thesis.

I cannot end without thanking my parents, my wife (Fauziah Yaman) and all my family for being so patient all these years I have worked on the Ph.D. Not only have they provided constant encouragement but also compromised on several ends to allow me to complete my thesis.

I would like to thank Directorate General of Resources for Science, Technology and Higher Education (DG-RSTHE) of Ministry of Research, Technology, and Higher Education of Indonesia for supporting financial expenses during my Ph.D. program.

Contents

論文の和文要旨	2
Abstract.....	4
Acknowledgements	5
Contents	i
1. Introduction	1
1.1 Background	1
1.2 Space Weather and Ionospheric Research	4
1.3 Nonlinear System Identification	6
1.4 Objectives.....	8
1.5 Outline.....	9
1.6 Significance.....	10
2. Radio Waves Propagating in the Ionosphere.....	12
2.1 Investigating the Ionosphere	13
2.1.1 Very Low Frequency (VLF) Measurement	13
2.1.1.1 VLF Wave Propagation in the Earth's Ionosphere Waveguide	14
2.1.1.2 VLF Transmitter	16
2.1.1.3 VLF Receiver	17
2.1.1.4 VLF Technical Architecture	19
2.1.2 Ionosonde.....	19
2.2 Solar Disturbances	21
2.2.1 Ionospheric Disturbances.....	22
2.2.2 Traveling Ionospheric Disturbances	22
2.3 Non-Solar Disturbances	23
2.3.1 Lightning -Induced Electron Precipitation (LEP).....	23
2.3.2 Early Events	23
2.3.3 Atmospheric Gravity Waves.....	24
3. Nonlinear Autoregressive with Exogenous Input Neural Network (NARXNN)	26
3.1 The Concept of an Artificial Neural Networks (ANNs)	26

3.2	Recurrent Neural Networks	27
3.3	NARXNN Structure	30
3.4	Training a Neural Network	32
3.4.1	Avoid Overfitting.....	32
3.4.2	Training Classification.....	34
3.4.2.1	Supervised training	35
3.4.2.2	Unsupervised training	35
3.4.2.3	Reinforcement training	35
3.4.3	Training Algorithm	35
3.4.3.1	Bayesian regulation training algorithm	36
3.4.3.2	Levenberg-Marquardt training algorithm	39
3.4.3.3	Scaled conjugate gradient training algorithm	43
3.5	Activation Function.....	46
4.	NARXNN Based Strategy for Prediction Modeling	48
4.1	General Strategy.....	48
4.2	Data Acquisition and Interpretation	50
4.2.1	Collecting the Data	51
4.2.1.1	The nighttime VLF electric field amplitude	51
4.2.1.2	Stratospheric temperature	53
4.2.1.3	foF2.....	54
4.2.1.4	Cosmic ray	54
4.2.1.5	Total column ozone	54
4.2.1.6	Dst index	55
4.2.1.7	AE index	55
4.2.1.8	Kp index	56
4.2.1.9	Mesospheric temperature.....	56
4.2.1.10	Solar radio flux at 10.7 cm index	56
4.2.1.11	Day of the year (DOY)	57
4.2.1.12	Sunspot number (SSN)	57
4.2.2	Data Cleaning	57
4.2.3	Data Average	58
4.2.4	Data Normalization.....	59
4.3	Training Strategy.....	59

4.3.1	Time Delay Selection.....	61
4.3.2	Hidden Layer Size	63
4.3.3	Training Algorithm	65
4.4	The Best Network Selection	66
4.5	Conclusions.....	68
5.	Applications for VLF Electric Field Amplitude Time Series Data.....	70
5.1	Inputs/output Data Processing	71
5.1.1	Low-Mid-Latitude VLF Propagation Path	71
5.1.2	Mid-Latitude VLF Propagation Path	75
5.1.3	High-Latitude VLF Propagation Path.....	78
5.2	One-step Ahead Prediction (OSA).....	81
5.2.1	Architecture of NARXNN OSA Modeling	82
5.2.2	Prediction Results	86
5.2.2.1	Low-mid-latitude VLF propagation path	87
5.2.2.2	Mid-latitude VLF propagation path.....	96
5.2.2.3	High-latitude VLF propagation path	105
5.3	Multi-step Ahead Prediction (MSA).....	114
5.3.1	Architecture of NARXNN MSA Modeling.....	115
5.3.2	Prediction Results	117
5.3.2.1	Low-Mid-Latitude VLF propagation path.....	117
5.3.2.2	Mid-Latitude VLF propagation path	122
5.3.2.3	High-Latitude VLF propagation path	127
5.4	Conclusions.....	132
6.	Application for foF2 Data.....	134
6.1	Inputs/output Data Processing	135
6.2	Architecture of foF2 NARXNN Modeling	138
6.3	Prediction Results	141
6.3.1	Hourly foF2 Prediction	141
6.3.2	Daily foF2 Prediction.....	146
6.4	Conclusions.....	151
7.	Conclusions and Future Works.....	153
7.1	Conclusions.....	153

7.2	Future Works.....	156
References.....		157

CHAPTER 1

1. Introduction

1.1 Background

The Earth's atmosphere is divided into four layers based its temperature structure. The uppermost layers are the mesosphere and the thermosphere. The ionosphere is the region coexisting with mesosphere and thermosphere wherein a fraction of atmospheric species is ionized by high energy solar radiation. The studies made in this thesis deal with some of the important dynamical influences associated with the characteristic of electromagnetic radio wave propagation in the Earth's ionosphere cavity. Further, continuous wave propagation from low to high frequencies up to remote sensing such as ionosonde has been used to observe the ionospheric condition.

The terrestrial ionosphere is the electrically charge D-region of the upper atmosphere, reaching up to approximately 500 km of the atmosphere altitude. The structure of the ionosphere is layered, each with a different electron density. The lowest region is the D-region with an altitude of 60-90 km. Above D-region lies the E-region extending from 90 up to 150 km. The highest region is the F-region, with an altitude ranging from 150 km to 500 km.

Radiation emanating from the Sun incident on the Earth's atmosphere ionizes the uncharged atoms or molecules in the daytime hemisphere of the ionosphere, producing free electrons. Non-solar ionization sources range from precipitating energetic electrons to meteoric ionization and cosmic rays. These processes maintain the free electron concentration within the nighttime ionosphere.

The D-region is likely the least studied layer of the ionosphere, as it lies too high for balloons to probe and it is too low for in situ satellite measurements. Hence studying subionospheric propagating Very Low Frequency (VLF) radio waves with the frequency range from 3 – 30 kHz reflecting off the lower ionosphere is almost only the method to probe the D-region continuously [*Helliwell et al.*, 1973; *Inan and Carpenter*, 1987; *Dowden and Adams*, 1988].

Ionosphere which can be illustrated in the simple schematic picture (in Figure 1.1). It is customary to simplify the analysis by assuming that the upper boundary of the ionosphere is

a specular reflector which imparts a phase shift of either 0° or 180° on each reflection. This crude approximation is not justified since the ionosphere is an imperfectly conducting medium with anisotropic properties. At VLF (less than 20 kHz), *Wait* [1957] proposed that the ionosphere could be regarded as a sharply bounded ionized medium. Using wave matching techniques, he showed how the ionospherically reflected wave may be calculated in terms of the electron density, collision frequency, and the Earth's magnetic field.

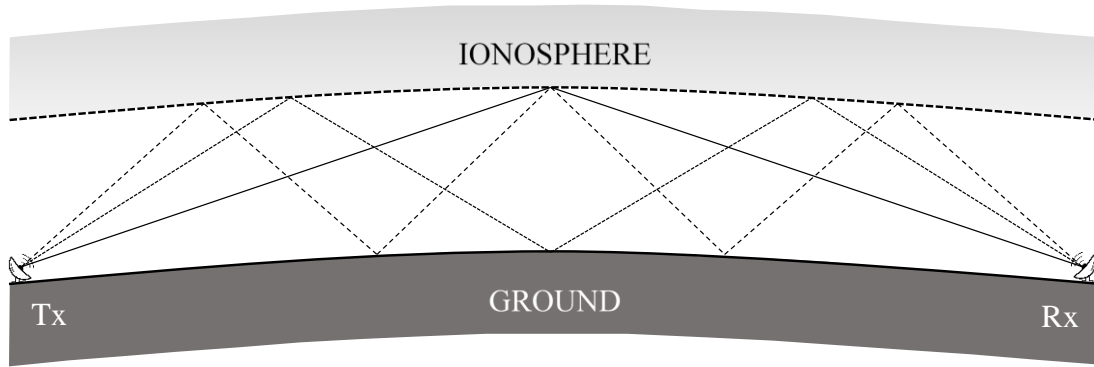


Figure 1.1: Wave-hop of VLF propagation paths. Tx and Rx are a transmitter and receiver respectively.

The VLF waves are a useful tool to study the lower ionospheric conditions because they largely reflect at the D-region of the Earth's ionosphere (60-90 km altitude). The D-region ionosphere perturbations have been observed in association with various geophysical sources as illustrated in Figure 1.2. Examples are both energetic particle precipitation and Transient Luminous Events (TLE's) due to thunderstorm activity [*Hobara et al.*, 2001; *Inan et al.*, 2010], solar eclipse events [*Clilverd et al.*, 2001; *Inui and Hobara*, 2014], gamma ray burst [*Inan et al.*, 1996; *Cohen et al.*, 2006], solar flares [*Schmitter*, 2013; *Singh et al.*, 2014], major seismic activity [*Hayakawa and Hobara*, 2010; *Rozhnoi et al.*, 2015], geomagnetic storms [*Peter et al.*, 2006; *Tatsuta et al.*, 2015], the impact of planetary waves [*Schmitter*, 2011; *Silber et al.*, 2016], mesospheric temperatures [*Silber et al.*, 2013] and solar zenith angle [*Thomson et al.*, 2014]. However, the temporal dependence of the VLF amplitude has both a complicated and large daily variability [*Clilverd et al.*, 1999; *Tomko and Hepner*, 2001] due to combinations of above mentioned different sources (both from below and above the ionosphere). Moreover, quantitative contributions from every external source are not yet known well. Furthermore, the highest ion density in the ionosphere is the F-region. Because in this region the plasma transport and chemical loss processes are balanced. The dominant atomic species is O^+ . The plasma processes are dominant in the topside region above the F-region peak where the primary ions

are O^+ and H^+ . Moreover, the F-region is further divided into two layers F1 and F2, the F1 layer is present during daytime whereas the F2 layer exists during the 24 hours on the day [McNamara, 1991]. The presence of the F2 layer typically affects propagation over a vast distance across the globe due to multiple ground and ionospheric reflections in the frequency range 2 to 16 MHz [Davies, 1989]. However, this region is susceptible to the growth of large-scale instabilities during night time. During such periods, depending on the evolution of instabilities, radio frequencies up to few GHz may get affected. The effect of the ionosphere on radio waves is frequency dependent and relates to the electron concentration in the ionosphere [McNamara, 1991].

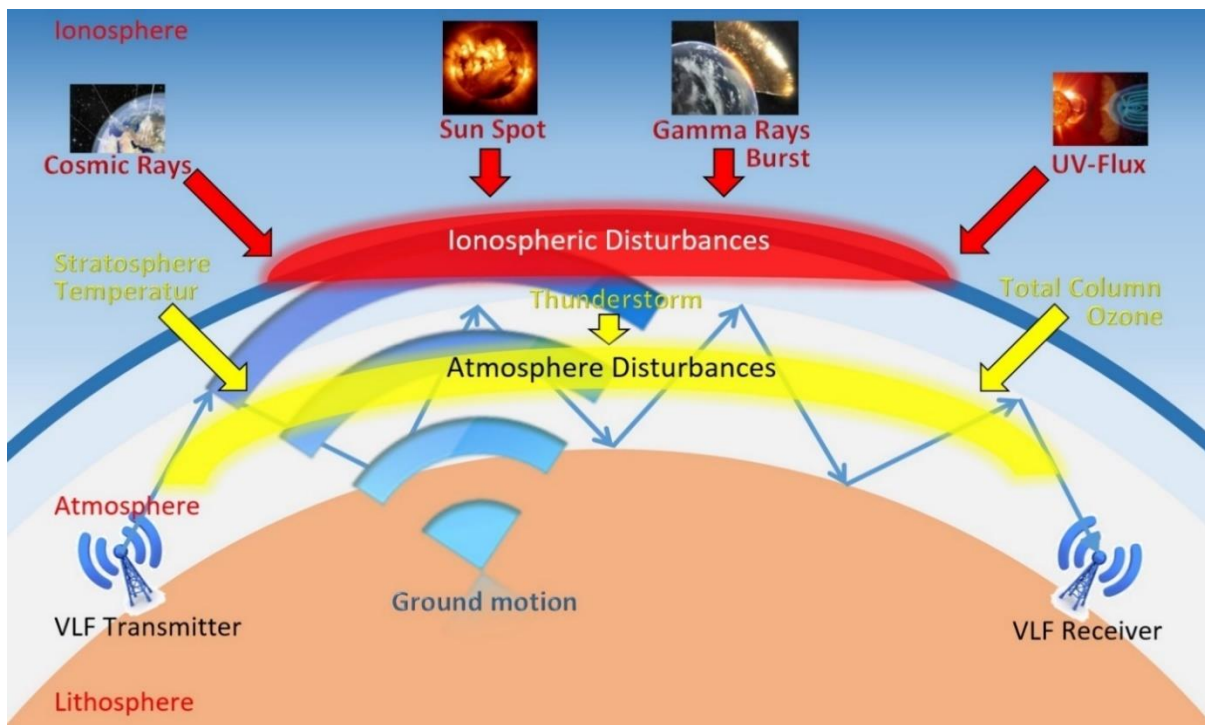


Figure 1.2 VLF propagation between the transmitter and receiver with perturbation along the path

The ionospheric sounding equipment called an ionosonde is one of the robust equipment to study the F2 region. Ionosonde is a radar which is capable of obtaining echoes from the ionosphere over a wide range of operating frequencies. Further, this measurement technique widely used to study the F2 region condition due to ionospheric perturbation such as perturbation neutral winds produce a part of storm-time changes in the electric fields through the ionospheric disturbance dynamo [Blanc and Richmond, 1980], the perturbation of the solar wind-magnetosphere dynamo [Senior and Blanc, 1984], the upper F-region perturbed under

the influence of the $E \times B$ drift [Hegai and Kim, 2016], the F2 region perturbation by ionospheric gravity and plasma pressure gradient current [Alken *et al.*, 2017], the peak F2-layer electron density change by following the Earthquake event [Hegai *et al.*, 2017], F-region of the ionosphere induced by atmospheric gravity waves (AGWs) generate solar eclipse [K. V. Kumar *et al.*, 2016], perturbation in the ionosphere from thunderstorm activity [K. V. Kumar *et al.*, 2016], complex perturbations during and after geomagnetic storms [Fejer *et al.*, 2017].

In this thesis, we use those two-measurement techniques to build a nonlinear dynamic model to predict the daily nighttime mean value of VLF electric field amplitude and critical frequency (foF2). Thus, this constructed model is an important issue to understand the lower-upper ionospheric responses from various external forcings and also detect the anomalies from unknown parameters.

1.2 Space Weather and Ionospheric Research

Space weather refers to the conditions on the Sun and in the solar wind, magnetosphere, ionosphere, and the thermosphere that can influence the performance reliability of space-borne and ground-based technological systems and can endanger human life or health [Haigh, 2007]. Space weather (Lyman-alpha, geomagnetic storm, and energetic particle precipitations) has a considerable effect on the satellites that run these technologies.

The particles from the solar wind leak through the magnetic field and interact with the Earth's magnetic field at lower altitudes where the Van Allen Radiation Belts are located, causing further changes to the Earth's magnetosphere [Horne *et al.*, 2007]. As the CMEs interact with the Earth's magnetic field, geomagnetic storms are induced. These storms are known to significantly affect radio wave propagation, as the effect of energetic particle precipitation during a magnetospheric shock event on VLF radio wave amplitude [M. A. Clilverd *et al.*, 2007]. The flare X-rays create significant extra ionization in the Earth's ionosphere, particularly in the lower D-region where VLF radio waves reflect thus lowering their reflection height [Thomson *et al.*, 2005]. The enhancement in the amplitude and phase of VLF signals by solar flares is due to the increase in the D-region electron density by the solar flare-produced extra ionization [A. Kumar and Kumar, 2014]. Furthermore, in the F-region, ionospheric storms change the neutral composition has been monitored by ionosonde measurement [Lastovicka, 2002]. The enhancement of the longitudinal conductivity gradient in the F-region induced by the energetic particle precipitation [Basu *et al.*, 2007]. Geomagnetic storms enhance the electron density in the F2 region [Burešová and Laštovička, 2007].

Moreover, it is well recognized that space weather induces severe ionospheric perturbations that can cause serious technological problems. Global position system (GPS) reference service is one of the communication systems which suffered because of ionospheric perturbations [Jakowski *et al.*, 2001]. Further, space flights and aviation system are particularly at risk due to ionospheric perturbation and satellite telecommunication problems [Akala *et al.*, 2012]

The ionosphere is defined as that portion of the atmosphere where free electrons and ions of thermal energy exist under the control of the gravity and magnetic field of the planet [Zolesi and Cander, 2014]. Figure 1.3 shows the vertical structure of the ionosphere illustrated by a representative ionospheric electron density profile for nighttime. The atmospheric altitude regions are defined based on atmospheric temperature also shown in Figure 1.3. The ionosphere is divided into four regions commonly by density structure the D, E, F1, and F2 regions. In the daytime, the solar Lyman-alpha (121.6nm) and X-ray radiation dominate the lower ionospheric region forming processes. In contrast, in the nighttime, both solar and non-solar sources maintain the smaller free electron and ion concentrations, such as cosmic rays, meteoric ionization, and precipitating energetic electrons [Hargreaves, 1992].

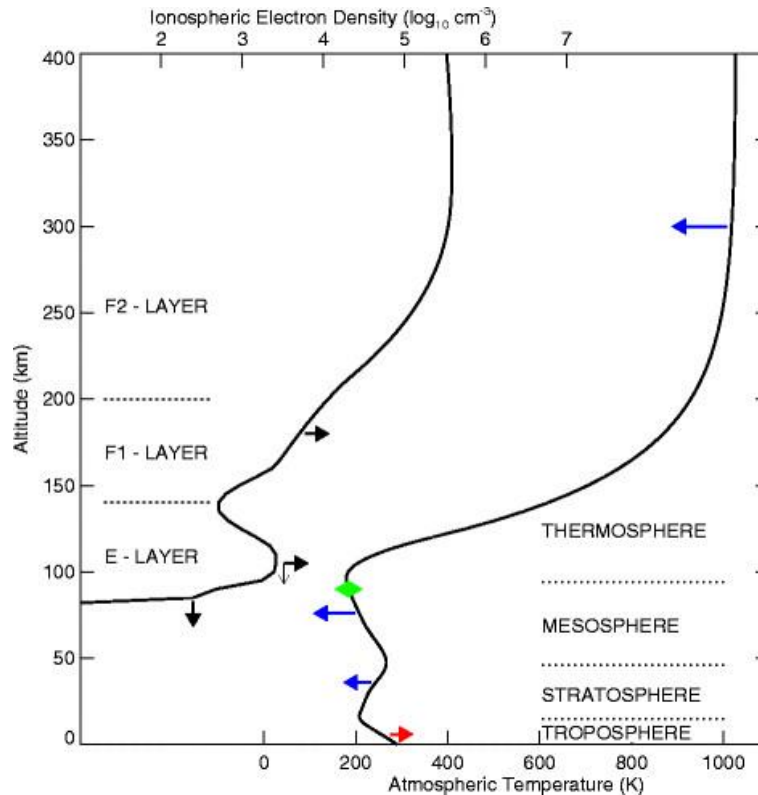


Figure 1.3: Typical temperature profile of the atmosphere and electron density profile of the ionosphere with altitude [Laštovička *et al.*, 2008].

The ionosphere is not as static the major component of space science is study of ionospheric dynamics through various means. The different techniques are capable of probing the different regions since the varying altitudes and electron densities of the ionospheric regions. Ionosonde is a swept-frequency pulse sounding, was the first technique employed and is still used today [Hargreaves, 1992]. The delay in the reflected signal under optimal conditions, as a function of frequency gives an almost direct measure of the electron density as a function of altitude. E- and F-region can probe by ionosonde using frequency from 1 to 20 MHz, transmitting modulated radio waves and receiving and analyzing the ionospherically reflected echo signal [Reinisch *et al.*, 1997]. However, a frequency lower than 500 kHz no echo is produced and need another measurement technique to investigate of the lower electron densities in the D-region [Rishbeth and Garriott, 1969].

The ionospheric measurement technique has been developed almost 40 years old called incoherent scatter radar (ISR) [Evans, 1969]. The electron density maximum of the ionosphere above the F2 region can be probe using this technique, and except electron density also able to measure other quantities, such as electron temperatures, ion composition, and electron [Evans, 1969]. The sophisticated signal processing, a large antenna, and a high-power transmitter are required for the measurement since incoherent scatter radar returns are rather weak, need a lot of sources also big facility area [Hargreaves, 1992].

The radio measurement techniques such as ISR and ionosonde are not capable to measurements of the D-region (~60–90 km) electron density, and balloons cannot reach this area because is too high also for the measurement of the satellite is too low. Using Langmuir probes in the rockets measurement can be used to measure in-situ electron density [Mechtly *et al.*, 1967]. The partial reflection technique is a method similar to the rocket technique, that the MF or HF signals transmit with a vertical incident from the ground station are reflected by irregularities conditions of the D-region also the profile of electron density in the medium can be obtained from the reflected signal characteristics [Belrose and Burke, 1964].

1.3 Nonlinear System Identification

System identification is a technique used to identify and characterize the mathematical relationship between the input and output of the system [Cullen *et al.*, 1996]. The mathematical model represents the relationship the input and output data based on the identification process [Ljung, 1999; Ogunfunmi, 2007]. The model should be capable of capturing all the dynamics of the system in a range of operating conditions. To proceed with a control engineering

technique in virtualized software systems, the linear and nonlinear dynamics should be characterized. Because of the lack of a first principle model of a software system, previous studies have started to investigate black-box models of a software performance management system in certain operating conditions [Hellerstein *et al.*, 2004; Åström and Wittenmark, 2008].

Traditional black-box models for nonlinear systems were based on the Volterra series [Giri and Bai, 2010]. This approach uses relevant integral kernels to describe the causal relationship between the system's input $u(t)$ and output $y(t)$. Wiener later used the Gram-Schmidt orthogonalization to develop another series expansion called Wiener series expansion [Hellerstein *et al.*, 2004], which expresses the system's output as a series of Wiener G-functional elements. Another way to interpret the Wiener model is in a block-structured manner which consists of a cascade connection between a linear dynamic system and a static nonlinearity using Hammerstein model in which the cascade connection order between the linear and nonlinear blocks is reversed [Hellerstein *et al.*, 2004].

In recent years, Artificial Neural Networks (ANN) has had an increasingly dominant impact on many engineering and scientific research areas, particularly in estimation purpose of such nonlinear mapping (representation between input and output of the system) [Arain *et al.*, 2012; Ayala and Coelho, 2016; De La Rosa *et al.*, 2017]. This approach has been well established as a universal approximation tool for nonlinear system fitting from input-output data, which is in the realization of an interconnected network. The network consists of several layers which operate in parallel (i.e. input layer, output layer and several hidden layers) and nodes (neurons where each node is connected to all nodes in the adjacent layers but not to the nodes in the same layer) to realize the nonlinear relationship of the input and output signals.

Most applications of the neural network in the field of space science are used to predict the geomagnetic indices or the ionospheric variations in F2 region. For example, prediction of the F2 region critical frequency (foF2) has been based upon feedforward neural networks [Nakamura *et al.*, 2009; Wichaiapanich *et al.*, 2017]. In addition, a neural network is used to forecast total electron content maps [Tulunay *et al.*, 2006; Ferreira *et al.*, 2017] and also used for predicting Dst index [Lundstedt, 2002; Revallo *et al.*, 2014]. Further, a neural network has been used to predict Kp index for the following 3h period [Boberg *et al.*, 2000]. However, simple NN using long-term temporal dependencies for learning purpose can be a difficult problem and less accurate [Lin *et al.*, 1998]. The problems of learning long-term dependency appear when the targeted output depends on inputs conferred at times far in the past and its difficult for the gradient-based algorithm. Because to avoid a degradation in the gradient produced by the partial derivative of the nonlinearity, the states do not need to propagate over

nonlinearities every time step. Further, nonlinear autoregressive with exogenous input neural network (NARXNN) is the useful method for overcoming the gradient problem in learning long-term dependency [DiPietro *et al.*, 2017]

The NARXNN is a hybrid of the dynamic recurrent neural network with nonlinear autoregressive with exogenous input model structure which allowing to use lagged and feedback from the output to the input regressor for the time steps in a finite number. Previously, the input regressor components composed of actual sample points of the time series and gradually took over by predicted values from previous values. The input regressor will start to be composed only for predicted values of the datasets. Moreover, the multi-step ahead prediction becomes a dynamic modeling task, which the neural network model operates as an autonomous system and attempting to recursively emulate the behavior of the dynamic system which has generated the nonlinear time series. DirRec strategy is one of the robust technique for multi-step ahead prediction. This strategy is a combination of recursive and direct strategies. DirRec strategy computes the predictions with different models for every time step, it enlarges the set of inputs by adding variable corresponding to the predictions of the previous step. However, the multi-step ahead prediction has a essential problem such as prediction error increase nonlinearly and external input will not available after several ahead prediction.

1.4 Objectives

The aim of this thesis is to develop a nonlinear model system identification based on the system identification technique for predicting ionospheric conditions through the radio wave propagation and remote sensing using external forcing parameters both from below and above the ionosphere and to study a quantitative contribution from external forcings of ionospheric properties.

The objectives of this thesis are shown as follows:

1. Developing a nonlinear system identification model of temporal ionospheric variations based on the combination between a neural network (NN) and a nonlinear autoregressive with exogenous input (NARX).
 - (a) Developing the best nonlinear autoregression with exogenous input neural network (NARXNN) model based on three different training algorithms, time delay selection and hidden layer size to achieve the highest Pearson's correlation coefficient (r), the lowest root meat squared error (RMSE), and the fastest computation time.

- (b) Developing a new technique by constructed hybrid NARXNN and DirRec strategy for long-term prediction to reduce prediction error.
2. Identifying a quantitative contribution from external forcings to study ionospheric property for science purpose.
 - (a) Analyzing the VLF electric field amplitude from three different latitude paths namely high-, mid-, and low-mid-latitude paths and two (Chofu and Tsuyama) VLF receiving stations to study the lower ionospheric (D-region) property.
 - (b) Analyzing the critical frequency in F2-region (f_oF_2) value from middle latitude Ionosonde station namely Kokubunji station to study the upper ionospheric (F-region) property.
 3. Forecasting the upper and lower ionospheric properties for engineering purpose.
 - (a) Predicting a one step ahead (OSA) of the daily nighttime D-region condition by VLF electric field amplitude in the different latitude paths and the hourly and daily F-region variability by f_oF_2 value.
 - (b) Predicting a multi-step ahead (MSA) of the daily nighttime D-region condition by VLF electric field amplitude in the different latitude paths.

The novelty of the thesis is applied for the first time a NARXNN technique in modeling VLF electric field amplitude time series. Then the temporal variability of VLF electric amplitude is predicted. Furthermore, the relative importance of various external forcings is identified to understand the physics of the daily D-region variability. Moreover, a new technique to reduce the error of multistep ahead prediction, which is a combination between the NARXNN and DirRec Strategy is developed. The novelty of the proposed multi-step ahead approaches is that all the input parameters used for the prediction are predicted values from NARXNN model and used in the DirRec strategy.

1.5 Outline

This thesis has been divided into 7 chapters. Chapter 1 will give a brief introduction to space weather and ionosphere. General knowledge of nonlinear system identification is explained.

Chapter 2 introduces the concept of radio wave propagation in the ionosphere. This chapter will give a knowledge of investigating the ionosphere using VLF radio propagation and sounding and also ionospheric variations due to different latitude.

Chapter 3 introduces the concept of artificial neural networks (ANNs) and gives a deep insight into all aspect of neural networks including the requirement to build a NARXNN model.

Chapter 4 is devoted to develop a nonlinear system identification using NARXNN. The data acquisition and interpretation of the inputs and output for the daily nighttime mean value of VLF electric field amplitude in mid-latitude path prediction model. The interconnection between the theoretical model structures and the resulting model used in the real application has been described as well.

Chapter 5 presents the results of one-step predictions of the nighttime mean values of VLF electric field amplitude using the NARXNN. Representing all process and developed model derived the input-output response in terms of past input and past output are described. Validation steps based on Root Mean Square Errors (RMSE) carried out to assess the performance of different models and data characteristics. Finally, the important input parameters are discussed in order to give the idea of external geophysical sources which influence the subionospheric VLF amplitude. Moreover, the extended prediction has been carried out in this chapter. Based upon the one-step ahead prediction modeling technique, the multi-step ahead prediction has been derived by incorporating the built structure model and new training method.

Based on the best nonlinear model in Chapter 4, Chapter 6 presents the results of foF2 prediction and also identify the physical external forcing influencing the critical frequency in the F2 region (foF2).

Chapter 7 presents the general conclusions of this thesis and also a recommendation for further studies.

1.6 Significance

The solution to the problem of a nonlinear system identification model for a quantitative contribution from external forcings to study and to predict the property of lower and upper ionosphere is developed by NARXNN model. Further, a method to overcome the problem of the prediction error increase nonlinearly is developed a new technique by combined NARXNN and DirRec strategy. The significance of this thesis is detailed as follows:

1. The NARXNN model using Levenberg-Marquardt algorithm neural network (LMANN) with three days delay time and two hundred neurons in the hidden layer to predict the

daily nighttime mean value of the VLF electric field amplitude variability in the three different latitude paths (high-, mid-, and low-mid-latitudes) is applied for the first time.

2. The new technique of hybrid NARXNN and DirRec strategy to reduce the multi-step ahead prediction error of the VLF electric field amplitude variability in the three different latitude paths (high-, mid-, and low-mid-latitudes) is developed.
3. The essentially physical contribution for the first time of lower (D-region) ionosphere by VLF electric field amplitude and upper (F-region) ionosphere by critical frequency in F2-region (foF2) value from various external forcings is identified.
4. The ionospheric variability of the lower and upper regions is predicted with the highest accuracy as shown by high Pearson's correlation coefficient and small root mean squared error.

CHAPTER 2

2. Radio Waves Propagating in the Ionosphere

There are two ways in which radio signals can propagate between a transmitter and receiver, at least one of which is on the Earth's surface. These are ground and sky waves propagation. Ground waves propagate along the curvature of the Earth while sky waves move to-and-from or through the ionosphere from either radar on the ground or onboard a satellite system.

In a case of sky waves propagating from a ground transmitter, the radio signal first encounters the D layer of the ionosphere where the electric field component of the ray forces the free electrons into oscillations at its same frequency. The charged particles then vibrate and collide with one another in-turn passing their energy to other particles in the process. Consequently, attenuation of the original signal from the radar takes place. The attenuation of the signal is inversely proportional to the radio wave's frequency squared. Hence, the transmitted lower frequency radio signals are more attenuated than their higher frequency counterparts.

The amount of signal loss is directly proportional to the number of particles present in the layer and its level of ionization. Therefore, only HF signals and those of higher frequency are able to propagate beyond the D layer. However, in both the E and F layers, HF signals show small attenuation in magnitude and significant refraction due to higher electron density concentrations. At some point in the ionosphere, this refraction becomes sufficient to send back the signals to the Earth's surface giving the impression that the rays have been reflected.

This reflection of signals depends on both the transmission frequency and incidence angle of the original ray. As the transmission frequency increases for a given angle of incidence, there comes a point where the maximum plasma frequency is exceeded and the signal propagates through the ionosphere into space. The angle at which the HF signals start making it through the ionosphere into further space without being refracted is called the Pedersen ray angle [Villain *et al.*, 1984].

2.1 Investigating the Ionosphere

2.1.1 Very Low Frequency (VLF) Measurement

Very low frequency (VLF) is radio frequencies (RF) with the range frequency of 3 to 30 kilohertz (kHz), corresponding to wavelengths from 100 to 10 km, respectively [Barr *et al.*, 2000]. The VLF band is used for secure military communication, a few radio navigation services and government time radio stations (broadcasting time signals to set radio clocks). The VLF waves are also used for military communication with submarines since VLF waves can penetrate at least 40 meters (120 ft) into saltwater.

The VLF radio waves can diffract around large obstacles and so are not blocked by mountain ranges or the horizon and can propagate as ground waves following the curvature of the Earth because of their large wavelengths. The main mode of long-distance propagation is an Earth-ionosphere waveguide mechanism [Hunsucker and Hargreaves, 2002]. The Earth is surrounded by a conductive layer of electrons and ions in the upper atmosphere at the bottom of the ionosphere called the D layer at 60 to 90 km (37 to 56 miles) altitude, which reflects VLF radio waves [Ghosh, 2002]. The conductive ionosphere and the Earth structures as a horizontal pipe which a few VLF wavelengths high and acts as a waveguide confining the waves so cannot possible to escape into space. The VLF waves travel in a zigzag path along the Earth, reflected alternately by the Earth and the ionosphere, in TM (transverse magnetic) mode.

VLF signal reflected in the daytime D-region mainly from the altitude range from 60 – 75 km, while in the nighttime, the electron densities are lower, and most of the reflection take place in the altitude range between 75 – 90 km. The electron densities take an important role for this reflection height because the density of the electron and hence refractive index increase rapidly in the space of wavelength with the altitude in this range, typically from a few per cm³ to several hundred or more per cm³ [Thomson *et al.*, 2007]. The electron density variability controlled by ionization from various external forcings. The solar Lyman-alpha is recognized as one of the variable that responsible for D-region ionization, which ionizes NO molecules. Solar radiation with the wavelength between 111.8 and 102.7 nm ionizes the excited O₂ molecules. All neutral molecules ionized by galactic cosmic rays and solar X-rays below 3 nm in D-region [Danilov, 1998].

VLF waves have a little of the fading experienced at higher frequencies and very low path attenuation, 2-3 dB per 1000 km [Hunsucker and Hargreaves, 2002]. The reason of these

conditions is the VLF waves reflected from the bottom of the ionosphere, while the higher frequency signals are returned to the Earth from higher layers in the ionosphere, the F-region, by a refraction process, and spend most of their journey in the ionosphere, so they are much more affected by ionization gradients and turbulence. Therefore, VLF transmissions are very stable and reliable and are used for long-distance communication. Propagation distances of 5,000 to 20,000 km have been realized [Ghosh, 2002]. However, atmospheric noise (sferics) is high in the band including such phenomena as a whistler, caused by lightning [Y Hobara *et al.*, 1995].

VLF waves at certain frequencies have been found to cause electron precipitation. VLF waves used to communicate with submarines have created an artificial bubble around the Earth that can protect it from solar flares and coronal mass ejections; this occurred through interaction with high-energy radiation particles [Clilverd *et al.*, 2009; Kolarski and Grubor, 2014].

2.1.1.1 VLF Wave Propagation in the Earth's Ionosphere Waveguide

Radio wave propagation at VLF has been studied experimentally and theoretically for past four decades. Several conceptual models are employed to explain the behavior of VLF radio propagation. For distances greater than a wavelength, where the near-field effects can be neglected, we can consider two primary methods of propagation: (1) ground wave and (2) sky wave. The total field can be considered as consisting of a ground and sky waves or sum as the sum of a number of modes in which electromagnetic energy is propagated between parallel boundaries as illustrated in Figure 2.1.

The Earth's surface, although quite heterogeneous in detail, can usually be considered as a plane sharp boundary layer with an effective conductivity which almost constant over the VLF band. In contrast, the actual ionospheric layer although apparently quite simply composed of charged particles with a varying height density gradient in a magnetic field is found to have electrical properties which vary greatly with time, frequency, geographic location, and direction of propagation.

Electromagnetic waves reflect when the incident upon conducting boundaries and guided along partially enclosed conducting structure. The surface of the Earth and the lower edge of the ionosphere act as good electrical conductors for VLF signals. The Earth's surface skin depth of seawater at 10 kHz is around 2.5 m and for dry Earth is around 500 m [U S Inan *et al.*, 2015]. While this skin depth is much less than the free wavelength of 30 km. The Earth-

ionosphere waveguide upper boundary consists of the lower ionosphere, which is a weakly ionized gas in which motion of ions and neutral molecules is often neglected [Budden, 1985]. The rate of electron collisions with the air molecules is determined by the electron and neutral temperatures [Chapman and Cowling, 1970]. The porous nature of the ionospheric boundary will also prevent the propagation of pure TE and TM modes. Instead, quasi-TE (QTE) and quasi-TM (QTM) modes propagate with a (typically small) field component in the direction of propagation. Additionally, these modes will propagate with different attenuations rates, propagation constants, and group velocities. Mode coupling will occur at sharp changes in boundary conductivity. Such discontinuities occur where the ionospheric conductivity changes rapidly with space (e.g., at the day/night terminator) or where the ground conductivity changes rapidly with space (e.g., land/sea interface). The curvature of the Earth also plays an important role because it will reduce the VLF signal amplitude on the order of a few dB.

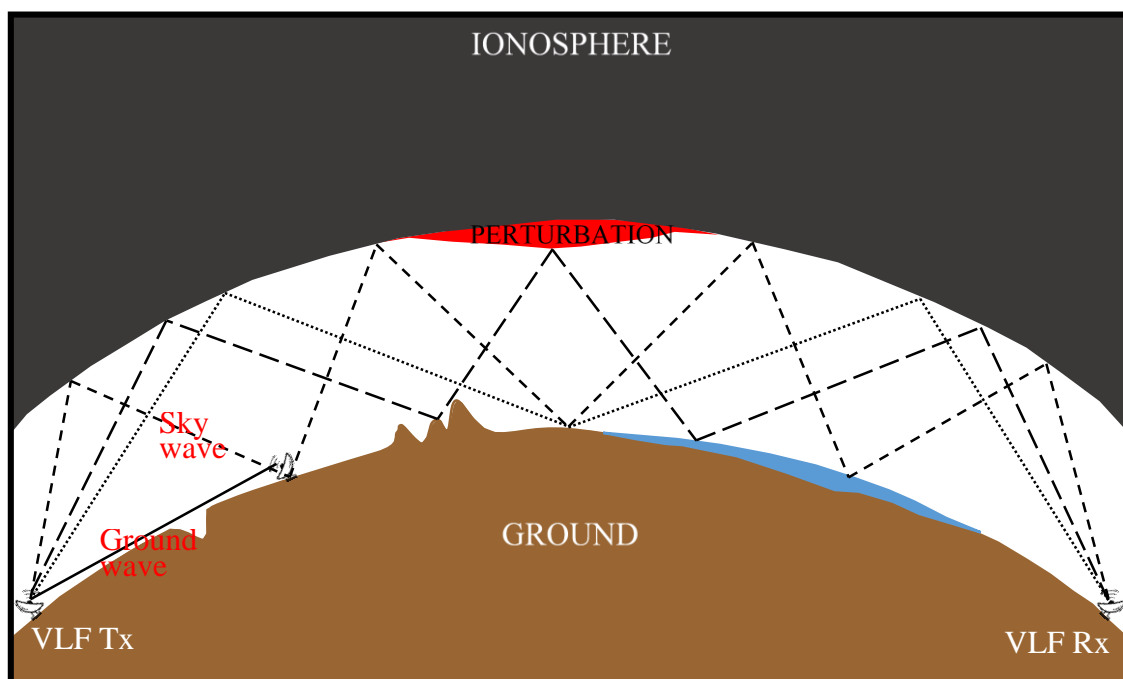


Figure 2.1: VLF wave propagation mode

Despite the complications of a more realistic Earth-ionosphere waveguide, however, the fundamental problem remains the same. The VLF signal will propagate in a multi-mode environment, and the signal detected at the receiver may simply be expressed as the sum of waveguide mode field values, each with different amplitudes and phases. It is noted at this point that the vast majority of numerical modeling methods for VLF propagation in the Earth-ionosphere waveguide output the sum-of-modes amplitude and phase (or the equivalent) as a

prediction. It is this sum of modes solution that the vast majority of VLF signal processing methods are intending to produce for comparison with theoretical modeling predictions.

2.1.1.2 VLF Transmitter

The high power VLF transmitter operated continuously by military continuously in order to communicate with its submarine force and research center for study the lower ionosphere. By utilizing the advantages of the Earth-ionosphere waveguide the transmitted signals are able to propagate very long distance with attenuation on the order of a few dB/m and penetrate sea water to a depth of several meters. Most of the transmitters use a 200-Hz bandwidth minimum shift keying (MSK) modulation. The bit sequence is projected to have a half Bernoulli distribution that is independent and identically distributed, giving an equal chance of receiving a 1 or 0 bit independent of the rest of the bit sequence. For each bit, the phase of the signal increases or decreases by 90 (or $\pi/2$ radians) over 5 ms [Gronemeyer and McBride, 1976], and the phase is forced to be continuous at bit transitions. Processing the MSK modulation is more complicated than processing a simple CW signal.

The MSK-modulated transmission can be described as [Forney, 1973]:

$$x_{MSK}(t) = \left(\omega_c t + \frac{\pi u_k}{2T} t + x_k \right), \quad kT \leq t \leq (k+1)T \quad (2.1)$$

where $x_{MSK}(t)$ is the transmitted signal, ω_c is the carrier or center frequency, u_k is bipolar data being transmitted at rate $R = 1/T$, and x_k is a phase constant which valid over the k th binary data interval $kT \leq t \leq (k+1)T$. Figure 2.2 shows the frequency shift keying (FSK) nature of the MSK waveform, with an upper frequency $\omega_c + \frac{\pi}{2T}$ being transmitted for $u_k = 1$ and lower frequency $\omega_c - \frac{\pi}{2T}$ being transmitted for $u_k = -1$. The tone spacing in MSK is one-half that employed in conventional orthogonal FSK modulation, giving rise to the name minimum shift keying. During each T second data interval, the value of x_k is a constant determined by the requirement that the phase of the waveform be continuous at the bit transition instants $t = kT$.

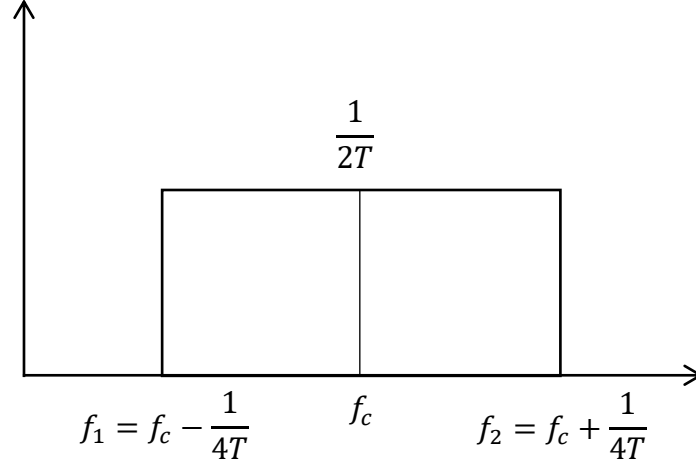


Figure 2.2: Tone spacing in MSK with the data rate of 1 bit per T seconds and carrier frequency of $\omega_c = 2\pi f_c$ rad/sec.

An abbreviated list of VLF transmitters which receive by UEC's VLF network is listed in Table 2.1. These transmitters are located in the United States, Japan, India, Australia. There are other VLF transmitters located around the world, but for the purposes of this thesis, the focus will primarily be on paths from these transmitters to receivers in UEC's VLF networks.

Table 2.1: A list of VLF transmitters received by UEC's VLF networks.

Sign	Location	Frequency
NLK	Jim Creek, Washington, USA	24.5 kHz
NPM	Lualualei, Hawaii, USA	21.4 kHz
NWC	North West Cape, Australia	19.8 kHz
JJI	Ebino, Miyazaki, Japan	22.2 kHz
JJY	Mount Ootakadoya, Fukushima, Japan	40.0 kHz

2.1.1.3 VLF Receiver

Understanding the received VLF signal (by analyzing the VLF receiver hardware) goes hand-in-hand with understanding the signal processing. Much of Stanford's work regarding receiver hardware has been published under the ELF/VLF Radiometer project [Fraser-Smith and Helliwell, 1985] and the purported AWESOME system [Cohen et al., 2010]. Other Universities have also contributed to the advance of VLF receiver hardware with the OmniPAL system from the University of Otago, New Zealand, [Dowden et al., 1998], the receivers used by the University of Washington for their World Wide Lightning Location

network [Lay *et al.*, 2004], and a wideband low-frequency receiver developed at the University of Bath, United Kingdom [Füllekrug, 2009]. In addition, softPAL VLF receiver developed by Dowden and Adams [2006] While receivers at the block level can be similar, there are variations between systems and it is important to understand the architectures used because design choices directly impact the performance of the system and the integrity of its associated data.

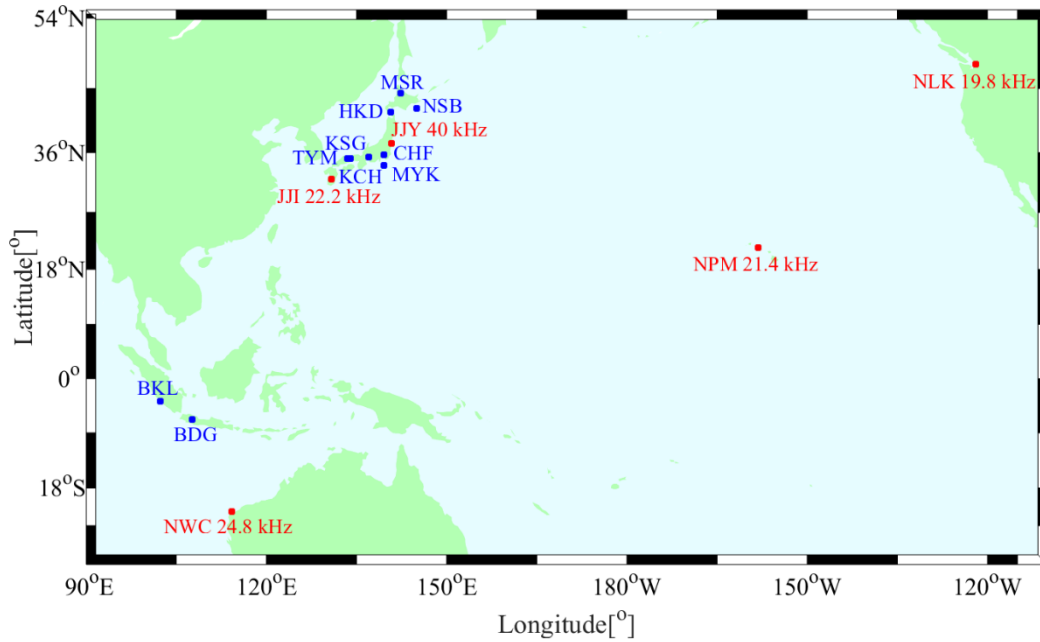


Figure 2.3: Geographical location of UEC's VLF networks. Red circles represent VLF transmitter and blue circles represent VLF receiver.

Receivers operated by University of Electro-Communications (UEC) Tokyo, Japan is located mostly in Japan and several stations in Indonesia. Figure 2.3 shows a map of currently operating VLF receivers and some important VLF transmitters. Different sites offer different background noise environments as well as different propagation paths from specific transmitters. Certain types of ionospheric events are more likely to be detected at different locations. Different paths also lead to different mode structures due to the different conductivity profiles produced by sea water, land, and ice. Using observations from multiple receiver locations also provides a diverse data set with which to analyze the spread spectrum processing method.

2.1.1.4 VLF Technical Architecture

SoftPAL is a PC based software VLF receiver which uses coherent detection and optimal demodulation of Minimum Shift Keying (MSK) and Interrupted Carrier Wave (ICW) signals to measure and record their phase (relative to GPS time) and amplitude. To measure the phase and amplitude of MSK signals SoftPAL uses an optimal center-frequency 2-bit demodulation algorithm, combined with modulation history subtraction (MHS) and lightning impulse (sferic) suppression. As a module running in LabChart for Windows, the SoftPAL system provides a full featured GUI for real-time display and analysis of the recorded data.

SoftPAL provides GPS locked timing and absolute phase measurements for signals at frequencies that are multiples of 1 Hz (SoftPAL can measure signals at non-integer frequencies but in this case, the measured phase will depend upon the time at which SoftPAL starts recording). By using sigma-delta Analog to Digital Converters (ADCs) and sophisticated digital signal processing, SoftPAL measures the time of a GPS 1 Pulse Per Second (PPS) signal relative to the ADC clock with an accuracy of a few nanoseconds (this is 1000 times less than the ADC output sample period) once every second. This timing difference is used to phase-lock a software frequency synthesizer to the GPS PPS. A typical modern GPS timing receiver has a timing error of ~ 25 ns. By using an oven crystal to generate the sampling clock for the ADC, the software phase-locked loop can average over 100 seconds worth of GPS PPS, achieving an overall timing accuracy that is significantly higher than that of the GPS receiver alone. The available VLF transmitters with adequate phase stability have either MSK (Minimum Shift Keyed) or ICW (Interrupted Continuous Wave) modulation. Other modulations (e.g. FSK) are not used by any phase stable VLF transmitters and are not supported. The SoftPAL VLF receiver can log several transmitters from up to three antennas at a time, logging phase and amplitude (PAL) with time resolutions ranging from 10 ms to 10 s. The number of antennas is limited to 3 because GPSNanoSync uses a 4 input ADC and one input is used by the GPS PPS signal.

2.1.2 Ionosonde

The ionosonde is an HF radar that is used to send radio energy pulses vertically into the ionosphere. The ionosonde radiates a modulated radio wave which is measured and analyzed if or when it is reflected by the ionosphere [Kunitsyn and Tereshchenko, 2003]. The receiver measures the time lag of the signal that is reflected from the ionosphere as a function of ray transmission frequency. Its output in simple terms is a graph of the time-of-flight of the

signal against the transmitted frequency. The reflection height is calculated under the assumption that the signal propagates at the speed of light c . In fact, the effective reflection height of the signal depends on the transmission frequency and the frequency-height plot is called an ionogram. An example ionogram is shown in Figure 2.4 below.

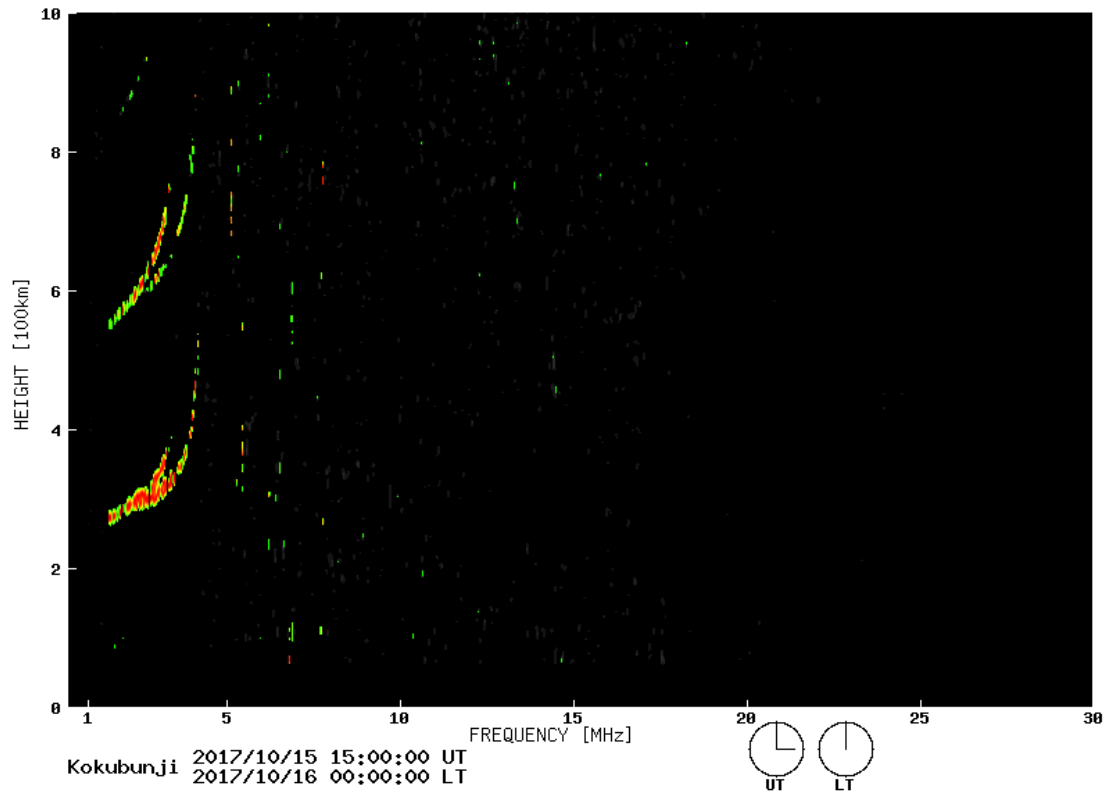


Figure 2.4: The plot shows an ionogram for Kokubunji, Tokyo, Japan (adapted from <http://wdc.nict.go.jp>).

The various ionospheric layers appear as approximate smooth curves separated by asymptotes at each layer's critical frequency. The layer's critical frequency and virtual heights are scaled from the asymptotes and lowest points on each curve respectively. The layer sections curve upwardly from their starting points due to the slowing down of the transmitted radio wave by the ionization with plasma frequency which is close to although not equal to the transmitted one. The red-like and green-like markings on the ionogram represents ordinary and extraordinary components of the radio wave propagating in the ionosphere. These two traces represent the virtual heights of the reflection points in relation to the plasma frequency of the constituent particles at each height or equivalently the transmitted frequency of the radio waves from the ionosonde.

An ionogram can be very complicated at times when there are activities such as Sporadic E, multiple hops of signals, D layer absorption, Spread F and Lacuna (gaps in the reflected radio signals ionosonde trace when turbulence occurs as a consequence of large electric fields causing the stratified nature of the ionosphere to be complicated).

2.2 Solar Disturbances

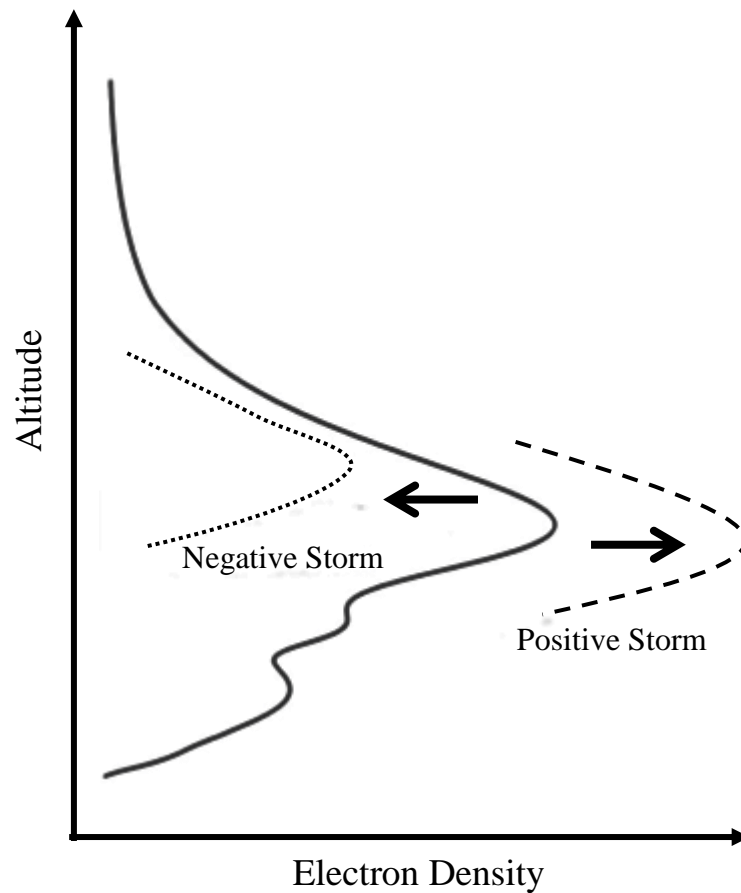


Figure 2.5: The vertical electron density profile change during a positive (dashed) and a negative (dotted) ionospheric storms in the F2 region, solid curve is the profile during normal ionosphere (adapted from [Hunsucker and Hargreaves, 2002]).

The solar wind energy increasing suddenly followed by increasing velocity and concentration of solar charged particle radiation or solar photon radiation in the upper atmosphere, lead electric currents and also heats in this region. Solar disturbances include geomagnetic activity and ionospheric disturbances caused increased Aurora intensity in the high latitude region. Ionospheric disturbances are an irregular and non-normal variation of the ionosphere and usually observed during geomagnetic activity. The solar disturbances affected

all regions in the ionosphere and F-region is the most perturbed especially around the peak of electron density.

2.2.1 Ionospheric Disturbances

In the middle latitudes, electron density could decrease or increase during the geomagnetic activity. To describe many disturbances which occur in the ionosphere associated with geomagnetic storm use ionospheric storm term. The ionosphere respond to the geomagnetic activity varies with season, time of day, longitude and latitude [*Dieminger et al.*, 1996]. The F2-region in the middle latitude region reacts to the geomagnetic storm in three conditions [*Hunsucker and Hargreaves*, 2002]. First, the initial condition is indicated by increased of electron density peak consider to pre-storm conditions continue up to a few hours on the first day of the storm. Second, the negative phase is represented by decreased of electron density to the condition before storm (pre-storm) occurred and can continue up to several days. Third, the recovery condition shown by the ionosphere condition gradually returns to the normal condition within the interval time from one to few days. Further, the vertical profile of electron density changing are shown in Figure 2.5.

The changing of the upper ionosphere composition, generate electric current and thermospheric winds could be produced by geomagnetic storm. The ionosphere structure in the normal condition influenced by those effects. The ionosphere electron concentration irregular changes caused by the variation in the transport process, recombination, and photoionization is can be made from one of the effects. The physics and the morphology of the storms in the ionosphere are not completely understood and still many interesting issues associate with these problems.

2.2.2 Traveling Ionospheric Disturbances

The traveling ionospheric disturbances (TIDs) is manifest of AGWs in the ionosphere as a result of collisional coupling between ionized and neutral particles. The ionosphere complications associated with the force direction of the neutral particles motion strongly modified by the geomagnetic field and the motion of particles limitation only within the magnetic field lines. The TIDs are divided into three categories [*Rieger and Leitingner*, 2002]: First, large scale TIDs (LSTIDs) are produced from geophysical events such as geomagnetic storm with altitude around 100 km or the signature of AGWs which occurred in the auroral region. LSTIDs have a horizontal phase velocity of 400-1000 m/s, periods ranging from 30

minutes to 3 hours and horizontal wavelengths greater than 1000 km. The propagation direction of the LSTIDs mostly from middle latitude to the equator region. Second, medium scale TIDs (MSTIDs) are produced mostly by AGWs in the lower atmosphere region. But, in the auroral region, the propagation of MSTIDs from middle latitude toward to equator can be the indication of AGWs. MSTIDs have a horizontal phase velocity of 100-300 m/s, periods ranging from 12 minutes to 1 hour and horizontal wavelengths from 100-1000 km. Third, small scale TIDs (SSTIDs) are produce only in the lower ionosphere region. SSTIDs have a horizontal phase velocity lower than 200 m/s, periods of a few minutes and horizontal wavelengths from 10-100 km. The LSTIDs occurrences are related with geomagnetic activity and not so many compare with MSTIDs which are commonly detected in the F2-region during the disturbed and quite ionosphere conditions in the daytime [Kalikhman, 1980]. The SSTIDs and MSTIDs are more complex movements compare with LSTIDs which mostly move toward to equator region.

2.3 Non-Solar Disturbances

2.3.1 Lightning -Induced Electron Precipitation (LEP)

Lightning-induced electron precipitation was identified by Michael Trimpi at Palmer Station, Antarctica in 1970's. This effect due to lightning activity can be detected by VLF signal measurement [Helliwell *et al.*, 1973]. The electromagnetic wave emanated by lightning couples to the magnetosphere interacts with energetic electrons in the Earth's radiation belts and scatters them onto the ionosphere below. Electron precipitation modifies the upper boundary of the Earth-ionosphere waveguide, resulting in a perturbation of the amplitude and phase of the received signal followed by a recovery lasting several 10's of seconds [U. S. Inan and Carpenter, 1987]. LEP research continues today in an effort to understand the mechanisms governing the dynamics of Earth's radiation belt population.

2.3.2 Early Events

Early events are produced in association with lightning. The time delay between the causative lightning return stroke and the onset of the event was much too short (< 100 ms) for the event to fit the physical mechanism for LEP [Armstrong, 1983]. Early VLF events have been related to transient luminous events (TLEs), although the specifics of the mechanism involved have yet to be satisfactorily identified [Moore *et al.*, 2003]. Further, the VLF receivers improved well over two decades and higher temporal resolution measurements were possible. This allowed researchers to identify between the relatively long delay of LEP events from the

parent lightning and other VLF events with a very short delay (<50 ms) from the parent lightning [Rodger, 1999].

2.3.3 Atmospheric Gravity Waves

The atmospheric waves are divided into three categories based on the origins and scale [Schunk and Nagy, 2009]. First, the largest scale waves include the planetary waves and atmospheric tides are propagate on global in nature and exhibit coherent patterns in both latitude and longitude. Second, smaller spatial scales are not global and related with the atmospheric gravity waves (AGWs) and the waves typically have a localized source and propagate with a limited range of wavelengths. Third, the smallest spatial scales are associated with acoustic waves. These waves, which are ordinary sound waves, do not play a prominent role in the dynamics or energetics of upper atmospheres.

The AGWs amplitude grows exponentially with the altitude in spite of maintaining a constant energy flux through the atmosphere which has decreasing density related to height (see Figure 2.6) [Matsushita and Campbell, 1967; Clark *et al.*, 1971]. The AGWs also can be produced between stratosphere and mesosphere region and continue propagate up to ionosphere and also possible generated from D- or E- regions. The AGWs mechanism to propagate from lower atmosphere to the ionosphere still not understood yet [Rieger and Leiting, 2002]. The earthquakes, volcanoes, the flow of air through the mountains, jetsreams are known as AGWs sources below the region between stratosphere and mesosphere, and Aurora in the high latitude region, the breaking of upward propagating tides, the movement of the solar terminator, and solar eclipses known as sources above the region between stratosphere and mesosphere [Afraimovich *et al.*, 1998].

Based on period and wavelength, AGWs are categorized into three [Hunsucker and Hargreaves, 2002]. First, large scale has horizontal velocities of 250-1000 m/s, wave period over an hour and horizontal wavelengths of about 1000 km. Second, medium scale has horizontal velocities of about 90-250 m/s, wave periods of about 15-70 minutes and horizontal wavelengths of few hundred km. Third, small scale has velocities less than 300 m/s, wave periods of 2-5 minutes and horizontal wavelengths smaller than AGWs in the medium scale.

The atmospheric gravity waves play an important role in the dynamics and energetics of the thermosphere, particularly in the altitude range from 100 to 250 km because AGWs interaction with the ionospheric plasma [Fesen *et al.*, 1995]. The F2-region is the important

region because electron density fluctuations and ionospheric plasma changes caused by AGWs propagation occurred mostly in this region.

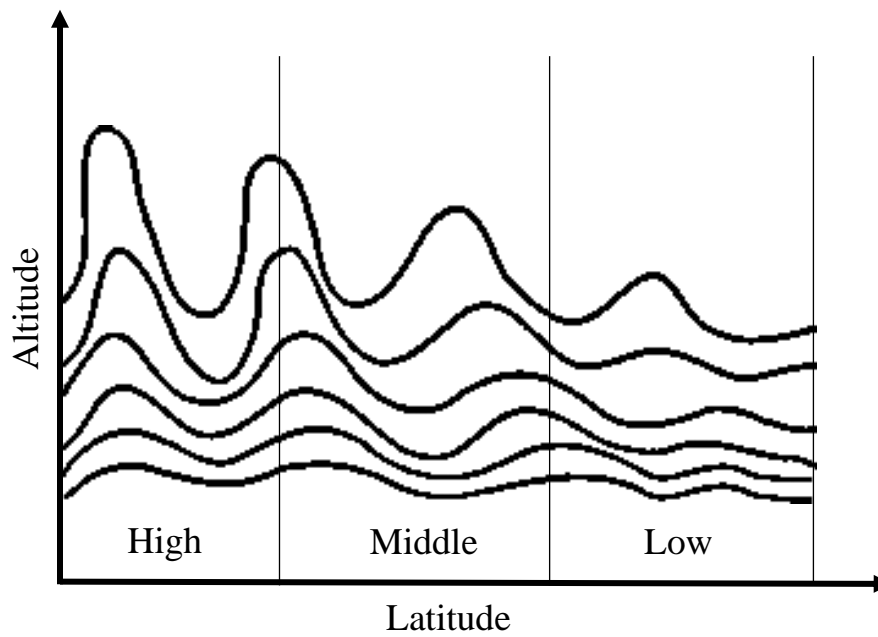


Figure 2.6: Gravity wave propagation characteristic in the different latitude from high to low regions (adapted from [Matsushita and Campbell, 1967]).

CHAPTER 3

3. Nonlinear Autoregressive with Exogenous Input Neural Network (NARXNN)

3.1 The Concept of an Artificial Neural Networks (ANNs)

The artificial neural networks, often abbreviated as a neural network, was invented by Warren S. McCulloch and Walter Pitts in 1943 [*McCulloch and Pitts*, 1943], which simulates the operations of the biological neural network. It has always been clearly understood that the mammalian brain functions in an entirely different manner from a digital computer and that the programming paradigms associated with modern computers do not replicate the information flow and decision-making processes that occur in the brain. The processing paradigm of the brain is believed to be one of massively parallel, non-linear, highly interconnected elements (neurons), with reconfigurable connectivity, able to re-organize its processing components to perform tasks and computation such as pattern recognition or perception many times faster than the equivalent function programmed in a digital computer. For example, the brain accomplishes perceptual recognition tasks such as identification of a familiar face in approximately 100-200ms. Estimates of the number of neurons in a human brain are approximately 10 billion, connected with 60 trillion synapses [*Haykin*, 2009]. The human learning process is believed to result from the development of the connections between neurons appropriate to matching examples of information provided by the senses.

Artificial Neural Networks (ANNs) are processing devices seeking to exploit a similar parallel, non-linear and configurable processing model as the mammalian brain, albeit on a very much smaller scale – a large ANN may consist of several thousand units, many orders of magnitude smaller (and still slower much slower) than the biological inspiration. Generally, an ANN implementation is not seeking to accurately resemble a biological system, but the concepts of learning by example and incrementally re-configuring the network such that it responds correctly are similar.

The applications of ANNs are today very wide, being routinely and commercially used for a variety of pattern recognition tasks including voice and optical character recognition and automated reading of vehicle registration plates. They are also extensively used in time series prediction in financial (market and sales predictions) and engineering fields. They find a use

for modeling of nonlinear control systems, for internet search tools and show promise in medical diagnosis.

3.2 Recurrent Neural Networks

One special form of neural networks is called recurrent neural network (RNN). The network can have single or multiple hidden layers of neurons. The fundamental difference between RNNs and feedforward neural networks is that they have one or more feedback loops. The feedback loop can appear in many forms between any two neurons or layers. Recurrent neural networks exhibit complex dynamics as they consist of a large number of feedforward and feedback connections [T. Chow and Cho, 2007]. These connections give them an extra advantage over feedforward NNs for handling time-series and dynamical related problems. A recurrent network with a smaller network size may be equivalent to a large or complicated type of feedforward NN architecture.

There has been a wide application of recurrent networks in the field of intelligent control, system identification and dynamical systems applications. In these applications, theoretical study of stability, the convergence of the network and their functional approximation capabilities are regarded as important. In [S. N. Huang *et al.*, 2005] an adaptive observer is proposed by using a generalized recurrent neural network where the learning is occurring on-line with no off-line learning. The overall adaptive observer scheme is shown to be uniformly ultimately bounded. In [T. W. S. Chow and Fang, 1998] the authors have presented a real-time learning control scheme for unknown nonlinear dynamical systems using recurrent neural networks (RNNs). A generalized real-time iterative learning algorithm is developed and used to train the RNN. The paper shows that an RNN using the real-time iterative learning algorithm can approximate any trajectory tracking to a very high degree of accuracy.

The use of recurrent neural networks as predictors and identifiers in nonlinear dynamical systems has received significant attention [Mandic and Chambers, 2001]. They can exhibit a wide range of dynamics, because of feedback, and are also tractable nonlinear maps. In this thesis, recurrent neural network models are considered as massively interconnected nonlinear filters with feedbacks that enable a more potential structural richness. The RNN architecture is used for prediction purposes, hence we present the material which is related to this aspect regarding the recurrent neural networks (Figure 3.1).

The basic building blocks of a discrete time predictor are adders, delays, multiplies and for the nonlinear case zero-memory nonlinearities. In order to use neural networks as nonlinear predictors, zero-memory nonlinearities such as threshold, piecewise-linear and logistics are required. These basic building blocks form the neurons in the NN architecture.

The inputs are assumed to be the delayed version of the neuron output ($y(k)$). In the problem of prediction, the nature of inputs to the network must capture some information about the time evolution of the discrete time signal or measurement.

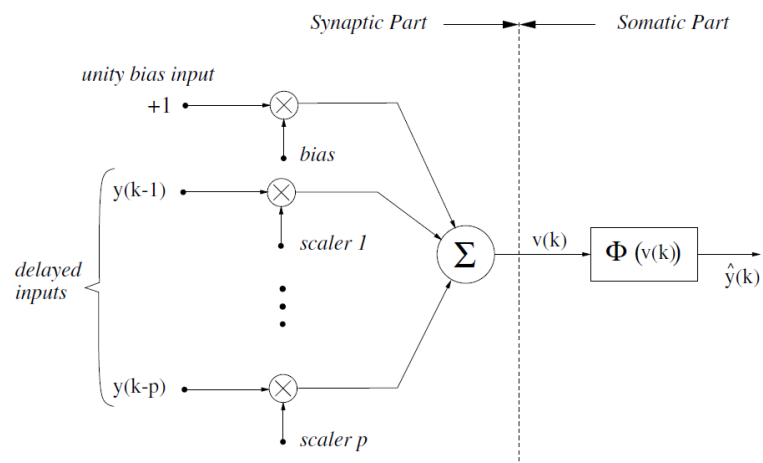


Figure 3.1: Structure of a neuron for prediction [Mandic and Chambers, 2001].

The simplest situation is for the inputs to be the time-delayed version of the output signal, i.e. $y(k - i), i = 1, 2, \dots, p$, which is known as a tapped delay line. This type of network provides a short-term memory of the signal. The overall predictor can be represented as:

$$\hat{y}(k) = \Phi(y(k - 1), y(k - 2), \dots, y(k - p)) \quad (3.1)$$

where Φ represents the nonlinear mapping of the neural network.

A typical recurrent neural network is depicted in Figure 3.2. If we consider connecting the delayed versions of the output of the network to its input, all together with the delays, one can introduce memory to the network and this structure becomes suitable for prediction. Information on the stability of such network can be found in [Mandic and Chambers, 2001].

As can be seen in the figure, the feedback within the network can be local or global. The global feedback is achieved by connecting the network output to the network input while

the local feedback is produced by introduction of feedback within the hidden layer. The output of an RNN, when used as a predictor, employing a global feedback, can be represented by,

$$\hat{y}(k) = \Phi(y(k-1), y(k-2), \dots, y(k-p), \hat{e}(k-1), \dots, \hat{e}(k-p)) \quad (3.2)$$

where $\hat{e}(k-j) = y(k-j) - \hat{y}(k-j)$. In other words, by adding the feedback and tapped-delay line to a static network architecture, we are adding memory to the network and then it becomes capable of prediction. This concept is shown in Figure 3.3.

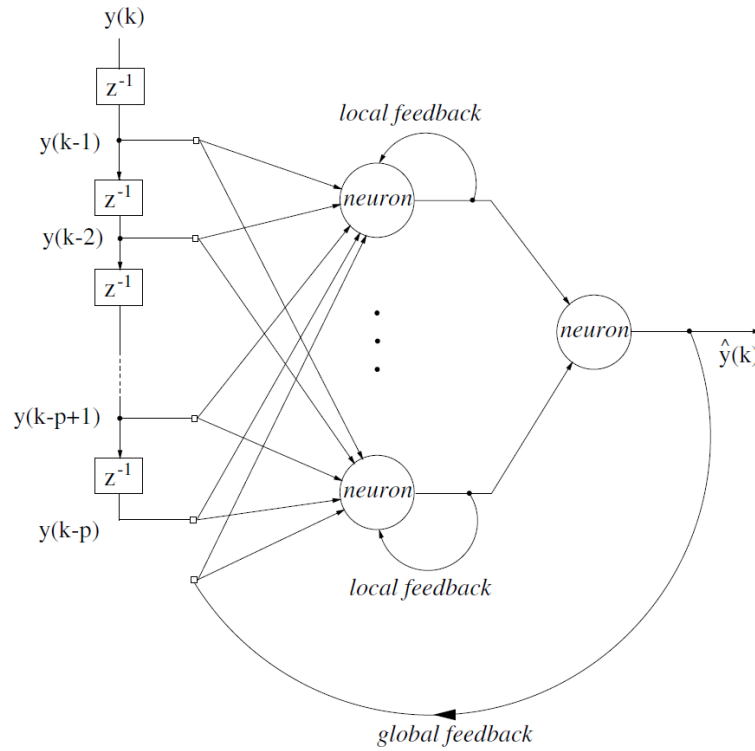


Figure 3.2: Structure of a recurrent neural network with local and global feedback [Mandic and Chambers, 2001].

The choice of the structure depends on the dynamics of the signal, learning algorithm, the prediction performance and the application. It is important to note that, there is no fast rule by which the best structure can be found for a particular problem [Tsoi and Back, 1997]. To introduce nonlinearity to the network, we use nonlinear activation functions. Any nonlinear function helps us achieve this goal, however for gradient-descent learning algorithm, this function $\sigma(\cdot)$ should be differentiable and belong to the class of sigmoid function. Surveys of neural transfer functions can be found in [Duch and Jankowski, 1999]. Examples of sigmoidal functions are:

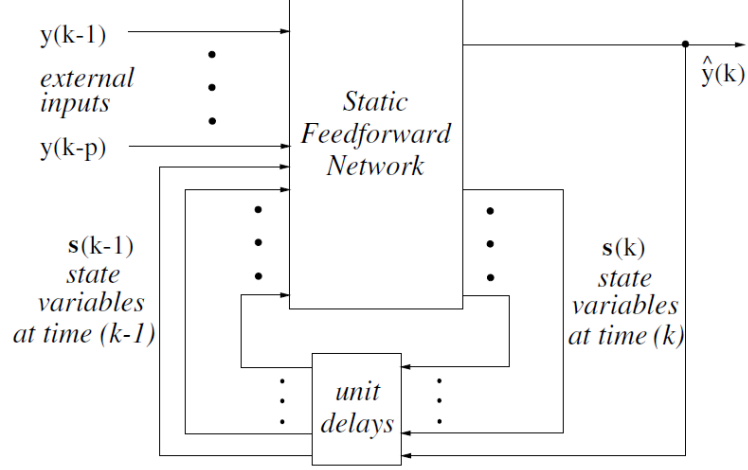


Figure 3.3: A canonical form of a recurrent neural network for prediction [Mandic and Chambers, 2001].

$$\sigma_1(x) = \frac{1}{1 + e^{-\beta x}} \quad (3.3)$$

$$\sigma_2(x) = \tanh(\beta x) = \frac{e^{\beta x} - e^{-\beta x}}{e^{\beta x} + e^{-\beta x}} \quad (3.4)$$

$$\sigma_3(x) = \frac{2}{\pi} \arctan\left(\frac{1}{2} \pi \beta x\right) \quad (3.5)$$

$$\sigma_4(x) = \frac{x^2}{1 + x^2} \operatorname{sgn}(x) \quad (3.6)$$

where β belongs to the set of all real numbers. According to Cybenko [1989], a neural network with a single hidden layer of neurons with sigmoidal functions and enough neurons can approximate an arbitrary continuous function.

3.3 NARXNN Structure

Nonlinear autoregressive with exogenous input neural network (NARXNN) is a combination of the recurrent and dynamic network as it has both feedback path and dynamical neurons or the tapped-delay line. NARXNN can be classified as a dynamic recurrent neural network that is capable of modeling efficiently time series with long-term dependences.

Feedback connections may enclose different layers of the network. The general architecture is shown in Figure 3.4. The model has a single input which is adapted to a tapped delay line memory of 2 units. It has a single output that is feedback to the input via another tapped-delay-line memory, also of 2 units. The contents of these two tapped-delay-line memories are used to feed the input layer of the multilayer perceptron. The present value of the model input is denoted by $u(k)$, and the corresponding value of the model output is denoted by $y(k + 1)$; that is, the output is ahead of the input by one-time unit. Thus, the signal vector applied to the input layer of the multilayer perceptron consists of a data window made up of the following components:

- ✓ present and past values of the input, namely, $u(k), u(k - 1), \dots, u(k - d_u + 1)$, which represent exogenous inputs originating from outside the network.
- ✓ delayed values of the output, namely, $y(k), y(k - 1), \dots, y(k - d_y + 1)$, on which the model output $y(k + 1)$ is regressed.

Figure 3.4 is referred to as a nonlinear autoregressive with exogenous inputs (NARX) model. The dynamic behavior of the NARX model is derived as follows:

$$\begin{aligned} y(k + 1) &= F[y(k), y(k - 1), \dots, y(k - d_y + 1); u(k), u(k - 1), \dots, u(k - d_u + 1)] \\ &= F[y(k); u(k)] \end{aligned} \quad (3.7)$$

where F is a nonlinear function, the next value of the dependent output signal $y(k)$ is regressed on previous values of the output signal and previous values of an independent (exogenous) input signal $u(k)$ at time k . The parameters d_u and d_y are memory delays. Two different architectures are defined for this type of neural network namely parallel and series-parallel architecture. For the parallel configuration the output regressor $y_p(k)$ is obtained as follows:

$$y_p(k) = [\hat{x}(k), \dots, \hat{x}(k - d_y + 1)] \quad (3.8)$$

where the P-mode contains d_y past values of the estimated time series. The second configuration is called series-parallel architecture in which the original output is used as a substitute of feeding back the estimated output.

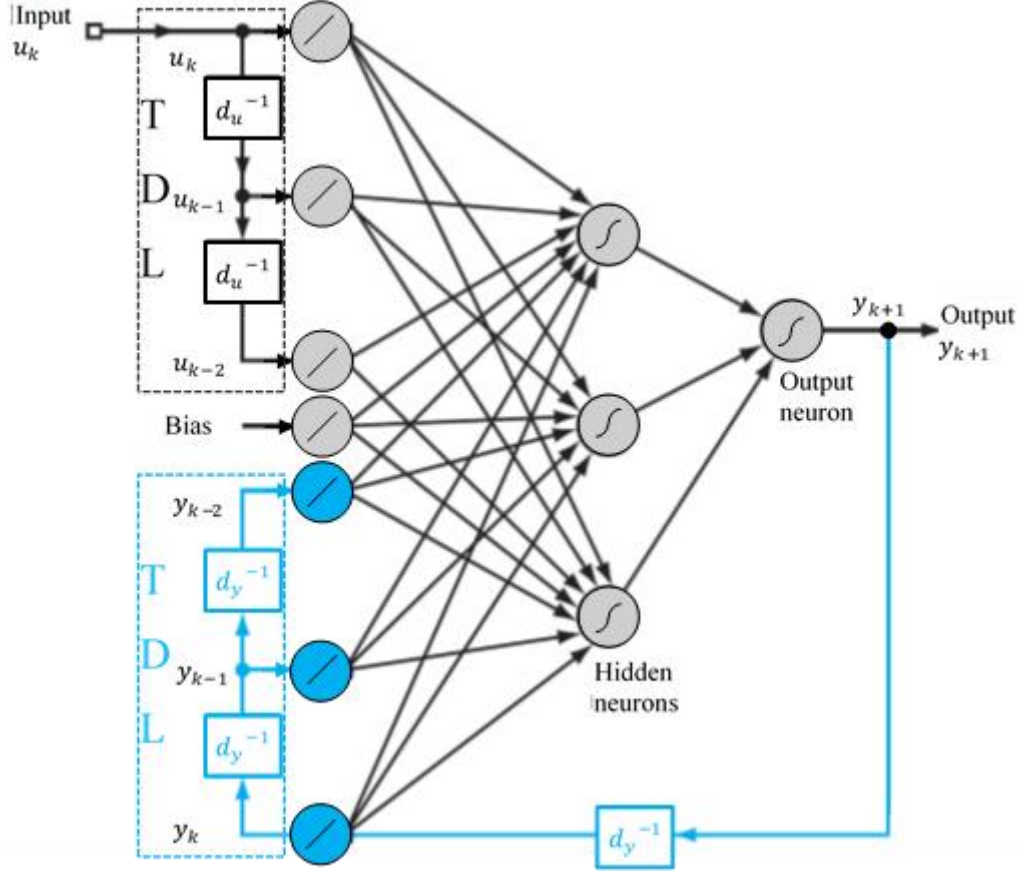


Figure 3.4: Nonlinear autoregressive with exogenous input neural network (NARXNN) architecture [Haykin, 2009]

Similar to the P-mode, in the SP-mode the output regressor $y_{sp}(k)$ is obtained as follows:

$$y_{sp}(k) = [x(k), \dots, x(k - d_y + 1)] \quad (3.9)$$

where the SP-mode contains d_y past values of the actual time series. This architecture has two advantages. Firstly, the input to the feedforward network is more accurate as it uses the past values of the actual time series not the estimated ones. Secondly the resulting network has a purely feedforward architecture, and static backpropagation can be used for training it.

3.4 Training a Neural Network

3.4.1 Avoid Overfitting

The important information which has to find for confirmation if the maximum generalization is reached. The case if the network maintain learning from the training set after

the maximum generalization point has been reached, it will break the related test set performance due to its overtraining. Moreover, when the validation error increases while the training error decrease known as overfitting. To avoid overfitting one of the solution is to reduce neurons number in the hidden layer [Pedroni *et al.*, 2010].

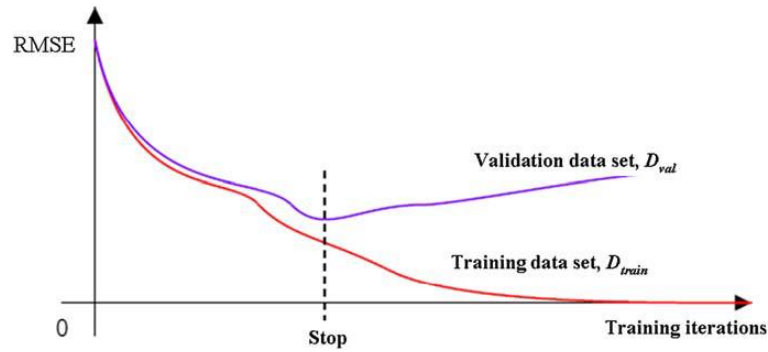


Figure 3.5: Overtraining example [Pedroni *et al.*, 2010].

The better approach to overcome problem in the training termination was presented by Palit and Popovic [2006]. The new technique developed by using pre-determine number of training steps for automatically stop the training. Moreover, the stopping strategy will stop the training process after the network has memorized all the problem in details and solve it. When, the network reaches the maximum generalization then the training process will stop at that stage. The minimum value is reached in this point which stopping the training process should be activated known as early stop technique. Except this condition, the network will perform in overfitting or overtraining condition.

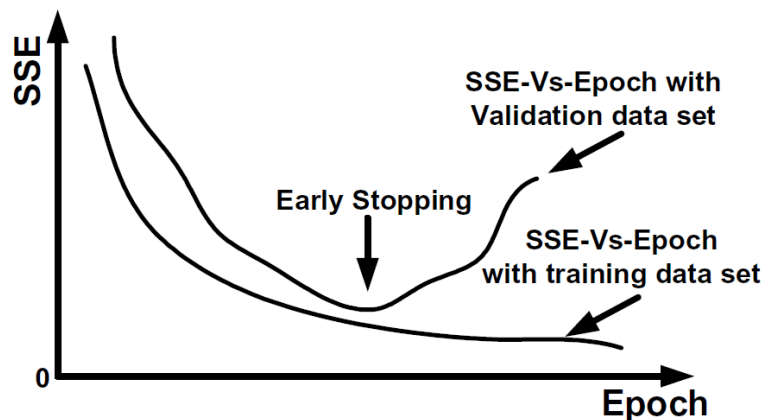


Figure 3.6: Early stopping of training [Palit and Popovic, 2006].

The cross-validation applied in early stopping is one of the techniques to prevent overfitting [Prechelt, 1998]. This technique applies the division of the dataset into a training dataset and validation dataset. Further, the training dataset also divided into training dataset, test dataset and validation dataset. The challenging problem up to now is to find the exact location of the early stopping. To overcome the stopping problem was developed technique in dividing the training set into the training error (e_{train}) represent average error per example across the training set, the test error (e_{test}) and the validation error (e_{val}).

The consequences of inappropriate training stop are underfitting and overfitting. Both of those problems should be prevented by the ability of the network to generalize the effect of the problem. In the overfitting problem, the network trained is very complex compared to the task to be learned. The network has to extract the structures of embedded noise and training set. In the underfitting problem, it is not so complex when training the network compared to the task to be learned. The large training dataset produce poorly identifies in the structures. The prediction results both of those conditions are not acceptable.

The network complexity is associated with the weights number and also determined by prediction accuracy of the constructed model. The overfitting and underfitting are associated with a degree of the function of fitting target and constrains to the network complexity. The deviation of the efficiency of learning network inside the training dataset concern with the generalization of the trained network. Moreover, the balance between overfitting which generates a large variance and underfitting which produces a high bias network is difficult to achieved.

3.4.2 Training Classification

The objective of training a neural network is generally to obtain a set of desired outputs from the network after the application of a set of inputs [Wasserman, 1989].

Training, therefore, entails presentation of a set of inputs, computation of error and adjustment of the weights as a result of this, to obtain convergence. This process continues until a certain specified training condition has been met. The three basic types of learning in the neural network are the supervised, unsupervised and reinforcement training [Engelbrecht, 2007].

3.4.2.1 Supervised training

In supervised training, the neural network is presented with a set of input and desired output vectors called a training set. The aim of this type of training is to present the network with a standard so that the network can adjust its weights until it is able to replicate this standard within a reasonable error limit. This is the most common type of training in neural networks [Engelbrecht, 2007]

3.4.2.2 Unsupervised training

In unsupervised training, there is no desired output vector or standard for the network to replicate; only input vectors are presented to the network. It is generally assumed that each input to the network arises from one of the several classes, and the network's output is an identification of the class to which its input belongs. Training of the network entails letting the network discover salient features to group the inputs into classes that it finds distinct [Engelbrecht, 2007]

3.4.2.3 Reinforcement training

This is a hybrid of the supervised and unsupervised training. It is not as commonly used as the other two. In reinforcement training, the desired outputs are not presented to the network but the network is allowed to know if it has replicated input signal well or otherwise [Engelbrecht, 2007].

3.4.3 Training Algorithm

The error backpropagation algorithm is one of the most significant technique for neural networks training processes developed from the steepest descent algorithm. The algorithm calculates the error for each example in the training dataset by using the difference between original and predicted output known as pre-defined error function. The error will propagate back over the hidden nodes to adjust inputs weights. When the minimum error solution has reached, the error back propagation procedure is completed. Still few limitations such as easily traps in local minima and slow convergence even though this technique id widely used in neural networks [Deyfus, 2005].

The small step sizes have to implement to not brake the required minima when the gradient is steep. The training process will be running very slow for a small constant time step

when the gradient is gentle. When the curvature of the error surface has different directions, the error valley could be occurred and it can be in slow convergence. The Gauss-Newton algorithm is capable to find adequate step sizes for each direction and could converge very fast by applied second-derivates of error function to evaluate the curvature of the error surface. Disadvantage of the second-derivates technique is the complexity of calculating process and cost much time.

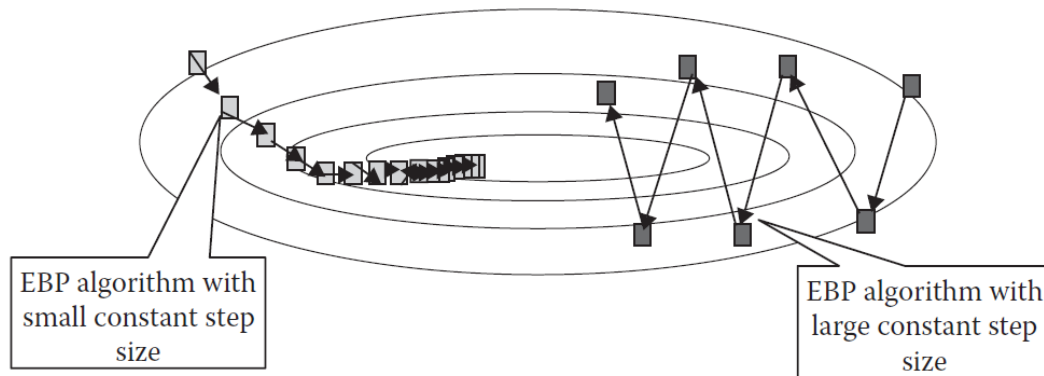


Figure 3.7: The steepest descent method with different learning constants. [Yu and Wilamowski, 2011].

The Levenberg-Marquardt algorithm is one of the technique to solve those problems and its suitable for small and medium sizes of the problems. Further, the Levenberg-Marquardt has a stable and fast convergence compared to the other methods. It combines the Gauss-Newton and the steepest descent algorithms. The speed of Gauss-Newton and more robust than the Gauss-Newton and also the stability of the steepest descent methods combine into the Levenberg-Marquardt. The idea is to combine both training processes for speed up the convergence the algorithm switches into the Gauss-Newton algorithm and for the point with the complex curvature to speed up the convergence, the algorithm approximately becomes the steepest descent algorithm until the local curvature is adequate to complete a quadratic approximation [Yu and Wilamowski, 2011].

3.4.3.1 Bayesian regulation training algorithm

Regularization techniques have been applied to neural networks with great success. Here is a list of the relevant work as it applies to the current research. *Burden and Winkler* [2009] states that Bayesian-regularized ANNs (BRANNs) are more robust than standard backpropagation nets, and can reduce the need for cross-validation. The cross-validation

process scales as $O(N^2)$, and therefore it is one of the most computationally intensive parts of using neural networks. BRANNs provide other advantages including lower overfitting probability and simpler networks. *Halpern* [2001] defines a measure of algebraic conditional plausibility represented using Bayesian networks. Uncertainty is represented using plausibility measures. The paper also shows the conditions under which conditional plausibility must be defined. The notion applies to both quantitative and qualitative Bayesian networks. *Sanghai et al.* [2005] introduce the concept of relational dynamic Bayesian networks, which are an extension of Bayesian networks to first-order logic. They are used for developing accurate fault diagnosis in factories, which are processes that have discrete variables with very large domains. The methods used by the authors are tested and shown to outperform standard particle filtering.

Further, Bayesian regularization with exponential weight decay in the sense that learning models that are well matched to the problem under study show a good correlation between generalization ability and the Bayesian evidence obtained [*MacKay*, 1992a]. Moreover, the Generalized Prediction Error (GPE) estimate of generalization performance [*Moody*, 1991]. The GPE index is based on the same author's work on the effective number of parameters in order to determine an optimal structure for the neural network. The idea is to balance the tendency of larger networks to overfit data with the value of the learning rate thus optimizing the generalization performance of the net. The author provides a method for calculating GPE in the case of quadratic weight decay [*Moody, J. E., Hanson, S. J., and Moody*, 1992].

Bayesian regularization is a training algorithm which updates the weights and bias values based on Levenberg-Marquardt optimization [*MacKay*, 1992b; *Dan Foresee and Hagan*, 1997]. The combination of squared errors and weights has to be minimized, and then the correct combination used to produce a network which generalizes well is determined [*Pan et al.*, 2013]. Bayesian regularization introduces the weights of the network into the training objective function denoted as $F(w)$ in equation (3.10) and further interpreted by *Yue et al* [2011].

$$F(w) = \alpha e_w + \beta e_D \quad (3.10)$$

with

$$\alpha = \frac{\gamma}{2e_w} \quad (3.11)$$

$$\beta = \frac{(N_D - \gamma)}{2e_D} \quad (3.12)$$

$$\gamma = N_p - \text{trace}(\mathbf{P}) \quad (3.13)$$

where e_w is the sum of squared error of the weights and e_D is the sum of network errors. And both α and β are the objective function parameters. γ represent for effective number of parameters, N_D is the number of data points, N_p is number of parameter, and \mathbf{P} is projection matrix of data onto the line of the best fit. The weights of the network are viewed as random variables in the Bayesian regularization framework, then the distribution of the training set and weights of the network are considered as Gaussian distribution. The α and β factors are defined by using the Bayes' rule. The Bayes' rule is presents two events (or variables) namely A and B , based on the marginal probabilities and conditional probabilities as shown in equation (3.14) [Li and Shi, 2012]:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (3.14)$$

where $P(A|B)$ is the conditional probability of A conditional on B , $P(B|A)$ the marginal of B conditional on A , and $P(B)$ the non-zero prior probability of event B , which functions as a normalizing constant. The objective function in equation (3.10) needs to be minimized to find optimal weight space, which is the equivalent of maximizing the conditional probability function given as in equation (3.15):

$$P(\alpha, \beta|D, M) = \frac{P(D|\alpha, \beta, M)P(\alpha, \beta|M)}{P(D|M)} \quad (3.15)$$

where α and β are the factors required be to optimized, D is the distribution of the weight, M is the particular neural network architecture, $P(D|\alpha, \beta, M)$ is the likelihood function of D for given α, β, M , $P(\alpha, \beta|M)$ is the uniform prior density for the regularization parameters, and $P(D|M)$ is the normalization factor. Maximizing the conditional function $P(\alpha, \beta|D, M)$ is equivalent of maximizing the likelihood function $P(D|\alpha, \beta, M)$. The result of this process, optimum values for α and β for a given weight space are found. Further, algorithm moves into Hessian matrix calculations take place and updates the weight space in order to minimize the objective function. Then, if the convergence is not met, algorithm estimates new values for α and β and the whole procedure repeats itself until convergence is reached [Yue et al., 2011]. The flowchart of BRANN calculation shows in Figure 3.8.

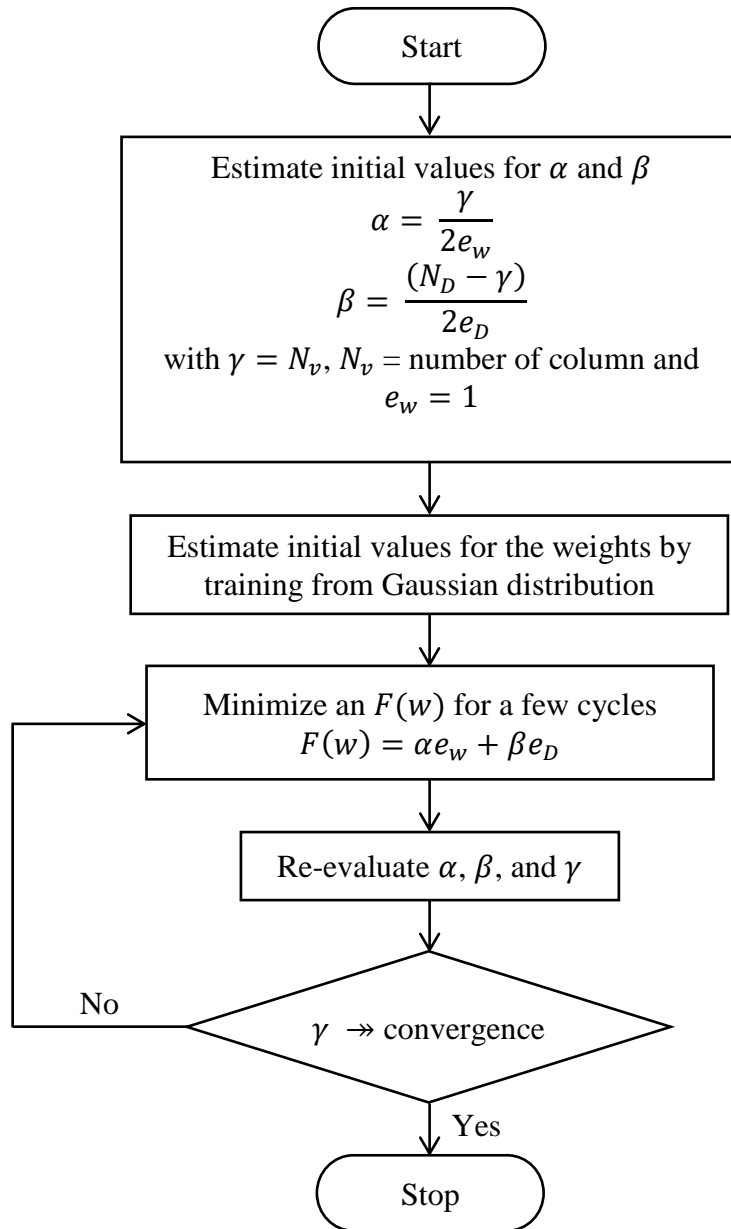


Figure 3.8. Flowchart of a BRANN algorithm.

3.4.3.2 Levenberg-Marquardt training algorithm

This is one of the learning procedures where the supervised learning problem is viewed as a numerical optimization problem. The error function is optimized with respect to the network weights. Using the Taylor series expansion we can express the error surface as [Haykin, 2009]

$$\varepsilon(w + \Delta w) = \varepsilon(w) + g^T \Delta w + \frac{1}{2} \Delta w^T H \Delta w + (\text{third} - \text{and higher} - \text{order terms}) \quad (3.16)$$

where g is the vector of the local gradient which is denoted by

$$g = \frac{\partial \varepsilon(w)}{\partial w} \quad (3.17)$$

The matrix H is the local Hessian which represent of curvature surface of the error performance and denoted by

$$H = \frac{\partial^2 \varepsilon(w)}{\partial w^2} \quad (3.18)$$

In the steepest descent method as represented by the back-propagation algorithm, the adjustment Δw applied to the network weights is defined by

$$\Delta w = -\eta g \quad (3.19)$$

where η is fixed learning rate parameter. In effect the steepest descent method operates on the basis of a linear approximation of the cost function in the local neighborhood of the operating point. This has the advantages that it is simple since it only uses first order effects. The inclusion of the momentum term in the update equation for weights is a crude attempt at using second-order information about the error surface. However, it makes the training process more delicate to manage by adding one more item to the list of parameters that have to be tuned by the designer.

Haykin [2009] states that in order to produce a significant improvement in the convergence performance on an MLP, higher-order information has to be used in the training process. A quadratic approximation of the error surface around the operating point can be employed to achieve that. Thus the optimum value of the adjustment applied to the weights is given by [*Haykin*, 2009]

$$\Delta w^* = H^{-1} g \quad (3.20)$$

where H^{-1} is the inverse of the Hessian matrix assuming that it exists. Equation (4.46) is the essence of Newton's method and it converges in one iteration if the error function is quadratic. However, *Haykin* [2009] states that the practical application of the Newton method is hampered by the following three factors:

- It requires calculation of the inverse Hessian, which can be computationally expensive.
- The Hessian matrix must be nonsingular. This is a very hard condition due to the nature of supervised learning problems which are often ill-conditioned.
- When the error function is non-quadratic, there is no guarantee that the Newton method will converge which makes it unsuitable for the training of a multilayer perceptron.

The Levenberg-Marquardt method, therefore, modifies equation (4.46) in such a way that it forces the supervised learning problem to become well-conditioned throughout the training process by the expression:

$$\Delta w^* = [H + \lambda I]^{-1} g \quad (3.21)$$

where I is the identity matrix of similar dimension to H and is the regularizing parameter that forces the sum matrix to be positive definite. If a multilayer perceptron with a single output neuron is considered, the error/cost function can be expressed as λ

$$\varepsilon(w) = \frac{1}{2N} \sum_{i=1}^N [y(i) - F(p(i); w)]^2 \quad (3.22)$$

where the training sample is given by $\{p(i); y(i)\}_{i=1}^N$ and $F(p(i); w)$ is the approximating function realised by the neural network and is given as a function of the inputs p at an instant i and weights w . The gradient and Hessian of the error function are given by

$$g = -\frac{1}{N} \sum_{i=1}^N [y(i) - F(p(i); w)] \times \frac{\partial F(p(i); w)}{\partial w} \quad (3.23)$$

$$= -\frac{1}{N} \sum_{i=1}^N [y(i) - F(p(i); w)] \times \frac{\partial F(p(i); w)}{\partial w^2} + \dots \quad (3.24)$$

$$= -\frac{1}{N} \sum_{i=1}^N \left(\frac{\partial F(p(i); w)}{\partial w} \right) \left(\frac{\partial F(p(i); w)}{\partial w} \right)^2$$

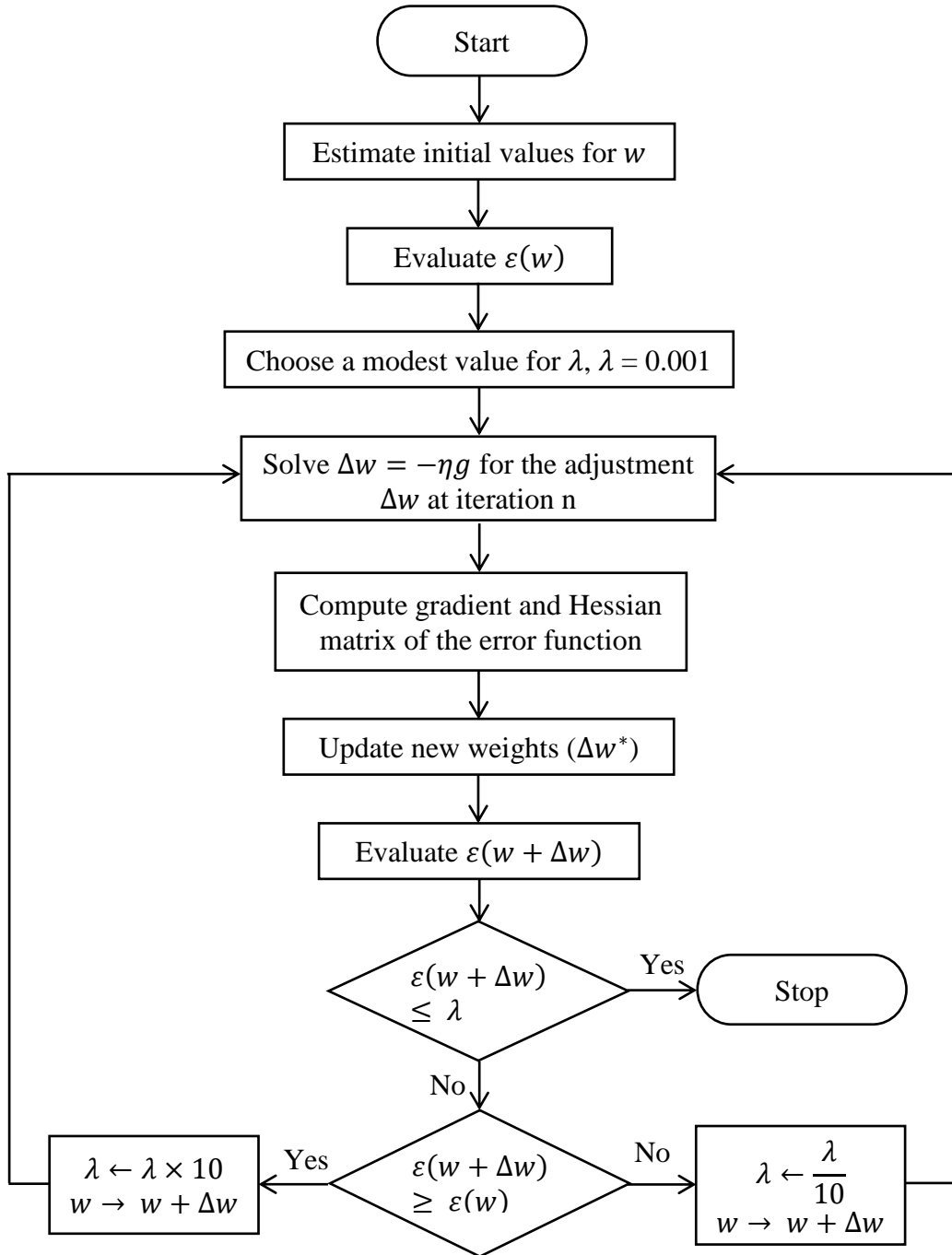


Figure 3.9. Flowchart of a LMANN algorithm.

These first-order partial derivatives can be efficiently computed using the back-propagation procedure. This second order partial derivative approximation is termed the outer product approximation of the Hessian matrix and its use is justified when the Levenberg-

Marquardt is operating in the neighborhood of the local or global minimum. The regularizing parameter λ plays a critical role in the way the Levenberg-Marquardt algorithm functions. If λ is assigned a value equal to zero, equation (3.21) reduces to a Newton formula; if a large value is assigned, the algorithm behaves like a gradient descent method. This switching functionality exploits the different capabilities of the Newton method and the gradient descent method. When the algorithm is operating near the optimum point, the Newton method causes it to converge rapidly towards the optimal point while the gradient descent method ensures that it does attain convergence through a proper selection of step-size parameter though at a very slow rate. The flowchart of LMANN calculation shows in Figure 3.9.

3.4.3.3 Scaled conjugate gradient training algorithm

The steepest descent negative direction of the gradient is the basic backpropagation algorithm adjusts the weights. The performance function of this direction is most rapidly decrease. This situation does not necessarily produce the fastest convergence although the function decreases most rapidly along the negative of the gradient [Hagan *et al.*, 2014]. In the conjugate gradient algorithms is produces generally faster convergence than steepest descent directions while preserving the error minimization achieved in all previous steps [Kişi and Uncuoğlu, 2005].

The learning rate of most of the training algorithms is used to determine the length of the weight update (step size). The step size is adjusted at each iteration for the conjugate gradient algorithms. The conjugate direction is a search made along the conjugate gradient direction to determine the step size, which minimizes the performance function along that line. The conjugate gradient algorithm start out by searching in the steepest descent direction at first iteration in equation (3.27) and used with line search. Further, to determine the optimal distance to move along the current search direction, the step size is approximated by a line search technique, avoiding the calculation of the Hessian matrix in equation (3.28). The next search direction is determined and conjugate to previous search direction in equation (3.31). The combination of the new steepest descent direction with the previous search direction become a general procedure for determining the new search direction [Hagan *et al.*, 2014].

The conjugate-gradient method based on the minimization of the quadratic function:

$$f(x) = \frac{1}{2}x^T Ax - b^T x + c \quad (3.25)$$

where x represent for a W by 1 parameter vector, A is a W by W symmetric, positive-definite matrix, b represent a W by 1 vector, and c is a scalar. The quadratic function $f(x)$ minimized by assigning to x the unique value:

$$x^* = A^{-1}b \quad (3.26)$$

Thus, minimizing $f(x)$ and solving the linear system of equation $Ax^* = b$ are equivalent problems.

The first iteration of the conjugate gradient algorithms start out by searching in the steepest descent negative direction of the gradient.

$$p_0 = r_0 = -g_0 \quad (3.27)$$

where r_0 represent for the initial residual as the steepest-descent direction and g_0 is the initial the gradient steepest-descent direction. Further, the optimal distance is determined by a line search which performed to move along the current search direction:

$$x_{k+1} = x_k + \alpha_k p_k \quad (3.28)$$

where

$$\alpha_k = \frac{p_k^T A e_k}{p_k^T A p_k} \quad (3.29)$$

$$e_k = x_k - x^* \quad (3.30)$$

α_k is learning rate parameter at step k and e_k represent for error vector at step k . Then, to determine the next search direction and conjugate to previous search directions. The combination of the new steepest descent direction with the previous search direction as a general procedure for determining the new search direction:

$$p_k = -g_k + \beta_k p_{k-1} \quad (3.31)$$

$$r_k = b - Ax_k \quad (3.32)$$

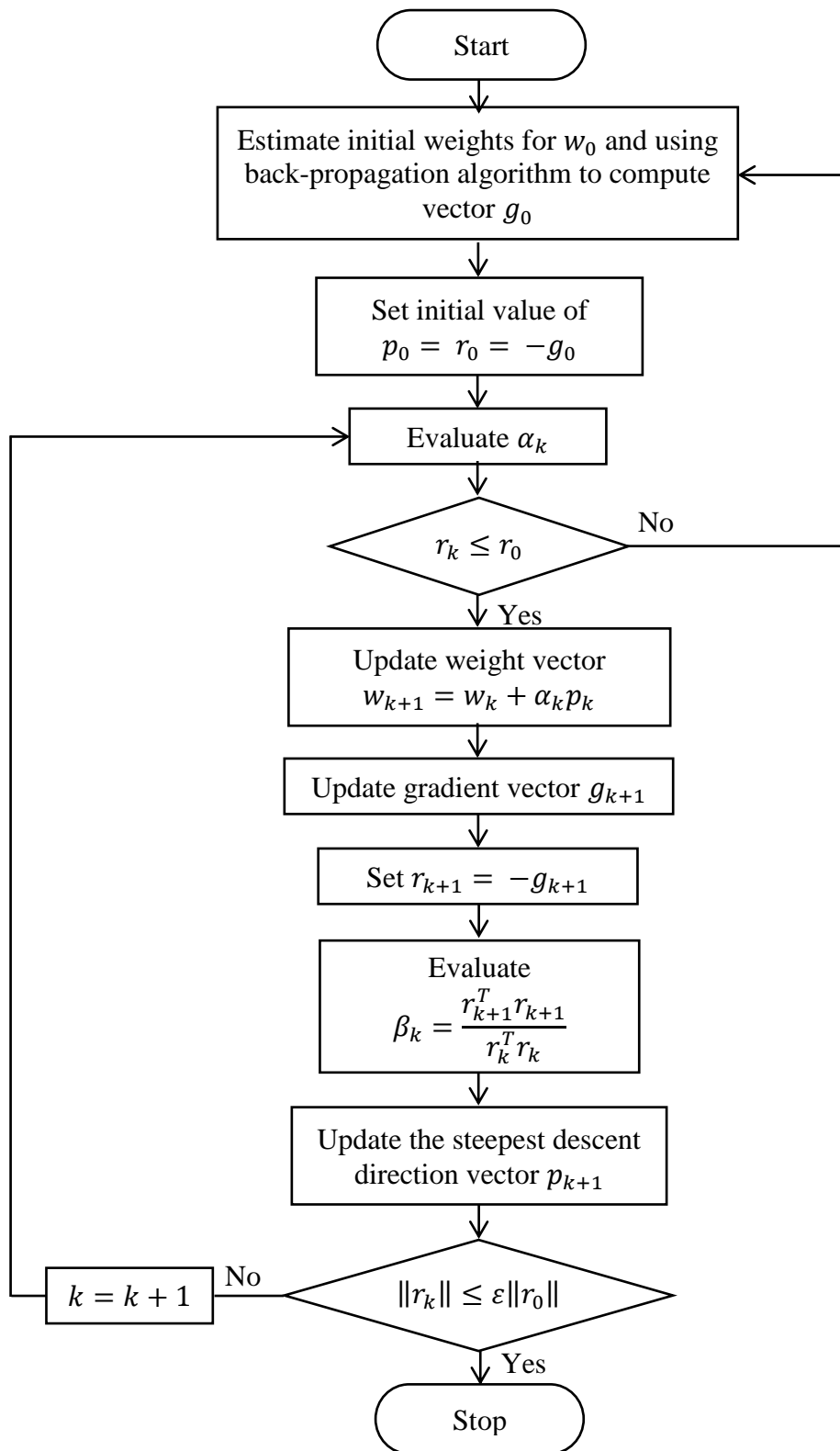


Figure 3.10. Flowchart of a SCG algorithm.

where β_k is a scaling factor at step k . The various technique of the conjugate gradient is well-known by the aspect in which the constant is computed. The procedure of the Fletcher-Reeves update as followed

$$\beta_k = \frac{r_k^T r_k}{r_{k-1}^T r_{k-1}} \quad (3.33)$$

The equation represents for the ratio of the norm squared of the current residual steepest-descent direction to the norm squared of the previous residual steepest-descent direction. Finally, to update the weight vector defined by

$$w_{k+1} = w_k + \alpha_k p_k \quad (3.34)$$

The flowchart of SCG calculation shows in Figure 3.10. Moreover, the advantages of the conjugate gradient algorithms are usually much faster than variable learning rate backpropagation, but the results will vary from one problem to another and require only a little more storage than the simpler algorithms. The conjugate gradient algorithms are usually a good choice for the networks with a large number of weights.

3.5 Activation Function

The ability of an artificial neural network to resolve complex non-linear engineering, scientific and social problems is largely dependent on the processing (hidden layer) unit activation functions. The activation function performs a mathematical operation to further process the net signal coming from the upstream neurons. Many activation functions are available for use in neural network implementation. We choose hyperbolic tangent or tangent sigmoid function as an activation function.

A hyperbolic tangent function maps all of its input to $[-1, +1]$. The shape of hyperbolic tangent function resembles a logistic curve (Figure 3.11). It has a limit of -1 when n approaches negative infinity, and a limit of 1 when n approaches infinity. The function and its derivative are:

$$f(x) = \frac{2}{1 + e^{-2x}} - 1 \quad (3.35)$$

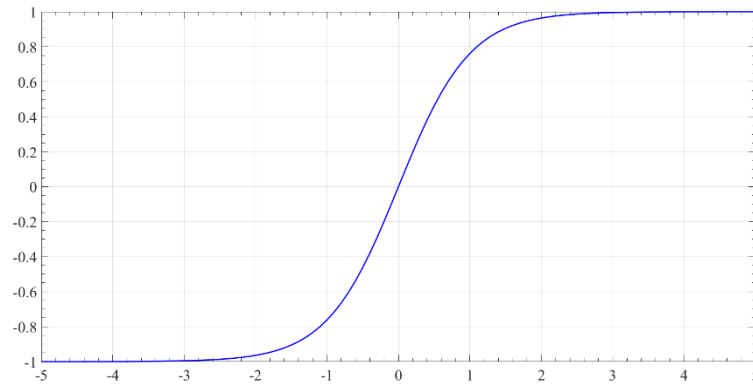


Figure 3.11: Hyperbolic tangent activation function

A hyperbolic tangent function performs better and faster training than the logistic function because of its symmetry (having positive and negative values). Therefore, it is ideal activation function of multi-layer perceptrons to be used in hidden layers.

It should be noted that the choice of activation function affects the network performance. The sigmoid function with a range of $[0,1]$ is the first popular to use in hidden layers while hyperbolic tangent with a range of $[-1,1]$ is in the second popular. However, hyperbolic tangent function tends to yield faster training because it produces both positive and negative values. In the output layer, the activation function should correspond to the distribution or property of the target values. For example, an exponential or absolute activation function is preferred if the target values are non-negative, while an identity activation function is selected for continuous-valued targets with no known bounds.

CHAPTER 4

4. NARXNN Based Strategy for Prediction Modeling

In order to model the dynamic changing of the ionospheric conditions in the lower and upper ionosphere region a general methodology is described in this chapter. This methodology is based on the knowledge of the external forcings which have influences in the dynamic changing of ion density in the ionosphere. Thus, first, the general strategy is described. This strategy is based on the knowledge of the parameters which have influences in the atmosphere and ionosphere region. Second, based on the general strategy, the NARXNN based methodology is described. This description is based on selections such as the NARXNN inputs and outputs, training strategy and the best network selection.

4.1 General Strategy

The objective of this study is to develop a nonlinear system identification which is capable of predicting ionospheric conditions through the radio wave propagation and remote sensing using external sources both from below and above the ionosphere. Thus, the parameters in Chapter 2 and Chapter 3 show that the

As said before, VLF waves propagate within the Earth-ionosphere waveguide and are extremely sensitive to perturbations occurring in the lower ionosphere region along their propagation path. Hence, measurements of these signals serve as an inexpensive remote sensing technique for probing the D-region. Further, VLF radio wave propagation cannot reach the upper ionosphere region so require a different technique to study the condition of this region. Ionosonde, one of the first radar sounding techniques, provide direct and accurate measurements of the ionospheric plasma density in the F-region. These two data measurements are represented the ionospheric conditions in the different latitude and altitude. Further, identifying physical parameters influencing in the lower and upper ionosphere and also different latitude region are important to study dynamic changing of the ionospheric conditions. Thus, the building of a physical model to predict VLF wave electric field amplitudes and foF2 are an important issue to understand the ionospheric responses from various external parameters and to also detect the anomalies from unknown sources

On the other hand, as shown in Chapter 4, the NARXNN is a suitable technique for modeling highly nonlinear systems like those represented in the ionosphere condition in Chapter 3. Besides, NARXNN is an excellent solution for modeling dynamic system (Section 4.3.4). Actually, one usual application of NARXNN is modeling dynamic system where the future events are predicted based on past or previous events. Furthermore, in this case, there are several data available to model ionosphere condition using VLF and ionosonde (foF2) measurement data. Thus, the strategy for NARXNN modeling of the dynamic changing of VLF and foF2 datasets are focused on giving (predicting) the real value of the complete dynamic with learning capability from the past datasets to generalize to new output conditions used during the training processes.

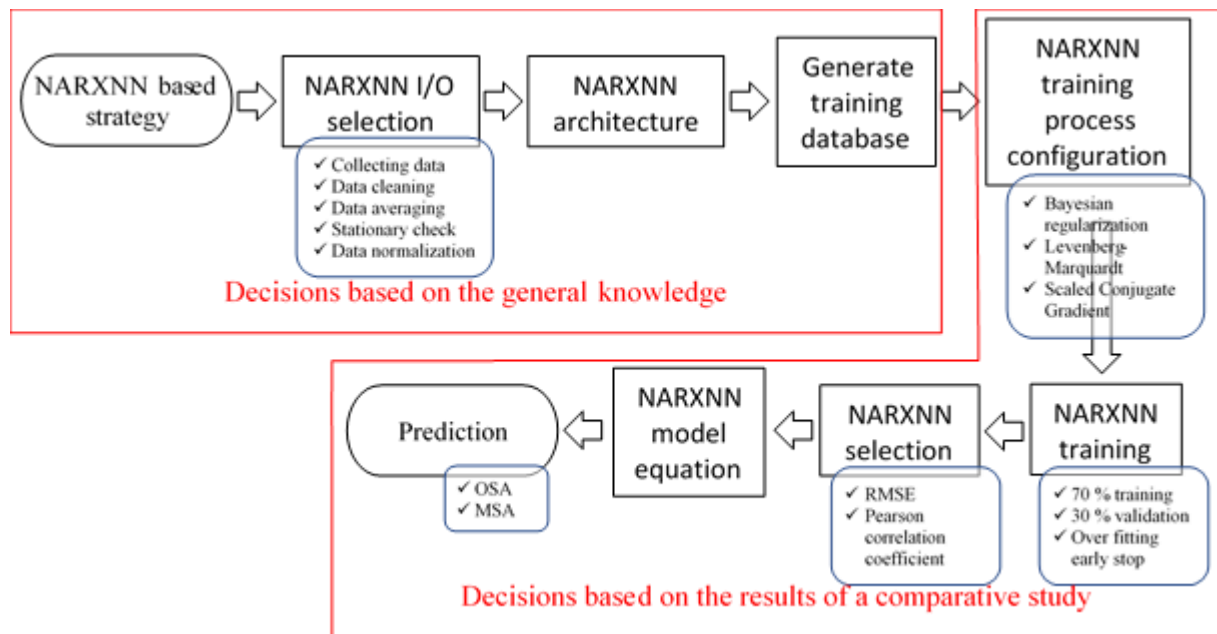


Figure 4.1: The NARXNN based strategy for modeling and predicting VLF and foF2 nonlinear dynamic measurements.

Therefore, as can be seen in Figure 4.1, the ANN strategy for predicting VLF and foF2 variables can be divided into two main sections. The first section is the configuration developed on the general knowledge of the process. Hence, the ANN input/output and architecture are decisions taken through the study from previous research, as well as restrictions such as external sources which have influences to the ionospheric conditions or build NARXNN architecture for modeling nonlinear dynamic datasets of VLF and foF2. However, other decisions like the training algorithm or the regularization method are taken after studying the state of the art.

On the other hand, the second section refers to decisions based on the results of a comparative study. For example, the network is trained with a different delay time in inputs/output, adjustment number of neurons in the hidden layer and compare the training results the appropriate number of neurons in the hidden layer is chosen. In fact, in most of the works related to NARXNN, parameters such as time delay for inputs and output, the number of neurons in the hidden layer, and the best training algorithm are selected with a trial-and-error approach.

4.2 Data Acquisition and Interpretation

In this section, we investigate the physical external datasets and data generation that was used In this thesis. NARXNN based methods are data-driven methods and the performances rely on the set of data they are provided with. The datasets that we use for training and testing our model are derived from many sources such as UEC's VLF network, NICT, Kyoto WDC, NASA, and Russian Academic of Sciences data center. The atmospheric and ionospheric parameters for the built model were described in Chapter 2 and Chapter 3 respectively. Furthermore, the target of this study divided into two main analysis, first for lower ionosphere used daily nighttime mean values of VLF electric field amplitude and second for upper ionosphere used hourly and daily mean values of foF2.

The selection of the input datasets was influenced by the physical studies carried out to understand the physical meaning of the relationship between VLF electric field amplitude changing and lower ionospheric condition in the different latitude paths. The data inputs datasets into the NARXNN were daily nighttime mean values of VLF electric field amplitude stratospheric temperature, cosmic rays, total column ozone, Dst, AE and Kp-indices, mesospheric temperatures, and F10.7 index. Further, we use nighttime values because VLF electric field amplitude on higher-order propagation modes appear in the Earth-ionosphere waveguide during the nighttime provide much larger variability in response to the change in the ionospheric condition than daytime. Therefore, the nighttime ionosphere is somewhat unstable and is thought to provide optimal condition to allow detection of comparatively weak ionospheric disturbances imposed on the VLF subionospheric propagation by various external forcing such as thunderstorms [Yasuhide Hobara *et al.*, 2001], magnetic storms [Tatsuta *et al.*, 2015], atmospheric changes [Silber *et al.*, 2013], Earthquakes [Hayakawa, 2007], and tsunami [A. Rozhnoi *et al.*, 2014]. The nighttime here is defined such that the whole GCP between transmitters and receiver in the completely night site. For low-mid-latitude corresponds to the

time interval from 15 UT to 19 UT. And for mid-latitude path corresponds to the time interval from 12 UT to 14 UT. Finally, for high-latitude corresponds to the time interval from 10:30 UT to 11:30 UT.

The selection of the input datasets was influenced by the physical studies carried out to understand the physical meaning of the relationship between foF2 measured value changing and F2 layer condition. The data inputs datasets into the NARXNN were hourly and daily values of foF2, DOY1, DOY2, F10.7, Dst, AE, and Kp indices, SSN. In addition, we use hourly and daily values of critical frequency (foF2) from ionosonde station in Kokubunji, Tokyo, Japan. Prediction accuracy of upper ionospheric conditions is critical for several applications affected by the space weather, including high frequency (HF) communications, satellite positioning, and navigation applications. This frequency range also sensitive to ionospheric storms, solar flare, and also protons emitted from the Sun.

4.2.1 Collecting the Data

For this study, the output of the prediction can be separated into two main sections. The first section is based on the data recorded by the VLF receiver located in Chofu (Japan). These VLF/LF receivers continuously measured the electric amplitude of signals from the three powerful transmitters such as NLK, NPM, and NWC. Moreover, the external parameter sources describing VLF amplitude in the nighttime such as stratospheric and mesospheric temperatures, cosmic rays, total column ozone, F10.7, Dst, AE and Kp-indices have been used to build accurate nonlinear modeling.

The second section, the ionospheric critical frequency, foF2, is one of the most important factors describing ionospheric conditions, and it is directly connected with the maximum usable frequency of HF communication links. This variable is recorded by the ionosonde located in Kokubunji (Japan). The long data for foF2 available in http://wdc.nict.go.jp/IONO/HP2009/ISDJ/manual_txt.html. Further, the external parameter sources which have influences with foF2 such as DOY1, DOY2, SSN, Dst, AE, and Kp indices are applied to construct a nonlinear modeling.

4.2.1.1 The nighttime VLF electric field amplitude

The daily nighttime mean electric field amplitude of VLF transmitter waves is the target variable for our prediction. The VLF transmitter signals from three powerful transmitters such as Hawaii, USA (NPM, 21.4 kHz, lat: 21.4°N, long: 158.1°W), Washington, USA (NLK,

24.8 kHz, lat: 48.2°N, long: 11.9°W), and North-West Cape, Australia (NWC, 19.8 kHz, lat: 21.8°S, long: 114.2°E) have been continuously monitored at The University of Electrocommunications (UEC) Tokyo, Chofu (CHF) with a geographic coordinate of latitude 35.6°N and longitude 139.5°E and Tsuyama, Japan (TYM, geographic latitude: 35.1°N, longitude: 133.9°E). This Great Circle Path (GCP) between the VLF transmitters and receiver is illustrated in Figure 4.2 and Figure 4.3 respectively. The UEC's network VLF receiving station is equipped with a vertical electric field antenna, SoftPAL VLF receiver unit, and the data logger (computer).

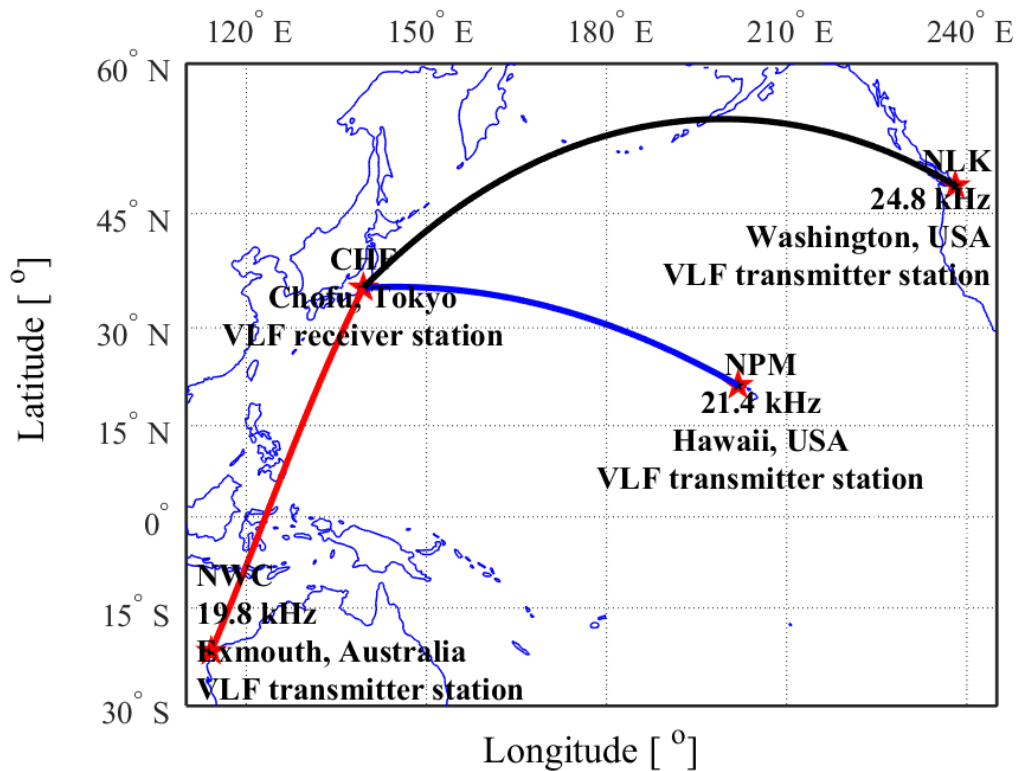


Figure 4.2: Geographical locations of the VLF transmitters and receiving site in Chofu (CHF) Tokyo, Japan. The solid curve indicates the Great Circle Path (GCP) between the transmitter and receiver (NLK-CHF-black curve, NPM-CHF-blue curve, and NWC-CHF-red curve).

In this thesis, we used the daily nighttime mean value of the VLF electric field amplitude data with a temporal resolution of the electric field amplitude data is 2-min because we are interested in time scale longer than few minutes. The nighttime here is defined such that the whole GCP between three paths respectively are nighttime condition corresponding to the time interval mid-latitude path from 12:00 UT to 14:00 UT, high-latitude path from 10:30 UT to 11:30 UT, and low-mid-latitude path from 15:00 UT to 19:00 UT at the receiving site (CHF). The VLF datasets were analyzed over the time interval from 1 January 2011 to 31 December 2013 for Chofu receiving station and 15 March 2014 to 26 May 2016 for Tsuyama station.

4.2.1.2 Stratospheric temperature

We used the daily nighttime mean of stratospheric temperature data in Kelvin at height of 30 km (± 1 km) was obtained from the Atmospheric Infrared Sounder (AIRS) level 3 data (<http://giovanni.gsfc.nasa.gov/giovanni/>). The AIRS sounding system is a suite of infrared and microwave instruments which carried on the NASA's Earth Observing System (EOS) Aqua satellite. The EOS Aqua platform is in a polar, sun-synchronous orbit that crosses the equator at 1:30 AM and 1:30 PM local time with an altitude 705 km, an inclination of 98.2° and an orbital period of 98.8 minutes. The AIRS instrument determines temperature with a spatial footprint is 1.1° in diameter, which corresponds to 15x15 km in the nadir and a vertical accuracy of 1°K per 1 km thick layer in the atmosphere [Aumann *et al.*, 2003]. This data is available in global observation but we only choose three paths namely NLK-CHF path (lat: 35°N - 48°N , long: 139°E - 121°W), NPM-CHF path (lat: 21°N - 35°N , long: 139°E - 159°W), and NWC-CHF path (lat: 21°S - 35°N , long: 114°E - 139°E).

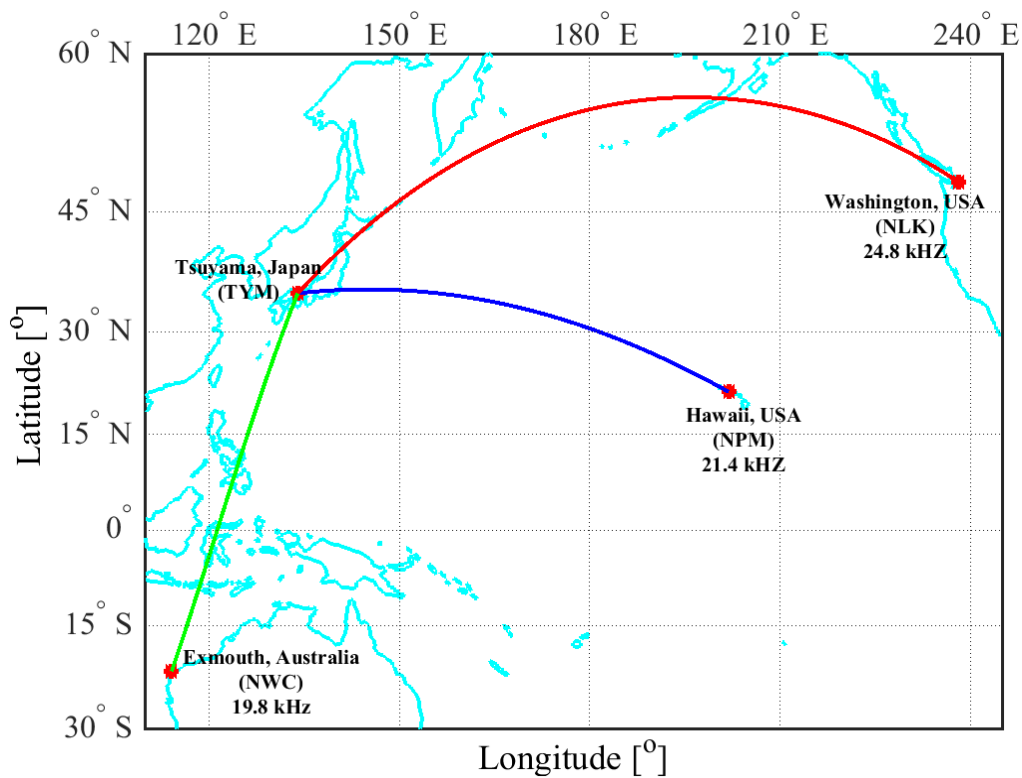


Figure 4.3: Geographical locations of the VLF transmitters and receiving site in Tsuyama (TYM), Japan. The solid curves indicate the Great Circle Paths (GCP) between the transmitters and receiver (NLK-TYM-red curve, NPM-TYM-blue curve, and NWC-TYM-green curve)

4.2.1.3 foF2

The F2 layer critical frequency (foF2), the maximum frequency that can be reflected by near vertical incident skywave (NVIS) of the ionosphere, is one of the most important parameters for planning HF communication systems and developing nonlinear ionospheric models. It has been extensively studied with ionosonde measurement. The data was obtained through the National Institute of Information and Communications Technology (NICT), a resource of the World Data Centre in Japan http://wdc.nict.go.jp/IONO/HP2009/ISDJ/manual_txt.html.

4.2.1.4 Cosmic ray

The nighttime daily mean of the cosmic ray data in count/min was obtained from Magadan cosmic ray station (60.04°N, 151.05°E) operated by the Institute of Cosmophysical Research and Radio Propagation Wave, part of the Russian Academic of Sciences (<http://cr0.izmiran.ru/mgdn/>). The count rate of the cosmic ray is the number of energetic particles originating from deep space that hit the atmosphere per minute per cm² of the active area of the sensor consisting of nine standard NM-64 proportional gas counter. The temporal resolution of the data is 10 s. The main source of ionizing radiation in the atmosphere at night is due to galactic cosmic rays. The ionization rate by the cosmic ray is related to the geomagnetic latitude due to the configuration of Earth's magnetic field [Störmer, 1930; Hofmann and Sauer, 1968]. The high intensity of the cosmic ray contributed to the sharpness change of electron density in the lower D-region [McRae and Thomson, 2004].

4.2.1.5 Total column ozone

The daily nighttime mean of Total Column Ozone (TCO) data in Dobson unit was obtained from the Ozone Monitoring Instrument (OMI) aboard on the NASA's AURA spacecraft (<http://disc.sci.gsfc.nasa.gov/Aura/>). The Aura spacecraft is a sun-synchronous polar orbit at altitude 705 km with an inclination of 98.2° and crosses the equator at 01:45 PM local time. The OMI is a backscatter spectrometer which measures the reflected solar radiation in the UV and visible part (270-500 nm) using three channels (UV-1 range 270-310 nm, UV-2 range 310-365 nm and visible range 365-500 nm) with a spectral resolution of ~0.5 nm. The angle view of the OMI telescope perpendicular to the flight direction is 114° which indicates the swath width of 2600 km on the Earth's surface with the spatial resolution of 13 km x 48 km for the UV-1 channel and 13 km x 24 km for the UV-2 and visible channels [Levelt et al.,

2006]. Daily global coverage is available from the data center but we take the limitation of mean values for three paths namely NLK-CHF path (lat: 35°N-48°N, long: 139°E-121°W), NPM-CHF path (lat: 21°N-35°N, long: 139°E-159°W), and NWC-CHF path (lat: 21°S-35°N, long: 114°E-139°E).

4.2.1.6 Dst index

The Dst or disturbance storm time index is a measure of geomagnetic activity used to assess the severity of magnetic storms. It is expressed in nanoteslas and is based on the mean value of the horizontal component of the Earth's magnetic field measured hourly at four near-equatorial geomagnetic observatories. Use of the Dst as an index of storm strength is possible because the strength of the surface magnetic field at low latitudes is inversely proportional to the energy content of the ring current, which increases during geomagnetic storms. In the case of a classic magnetic storm, the Dst shows a sudden rise, corresponding to the storm sudden commencement, and then decreases sharply as the ring current intensifies. Once the IMF turns northward again and the ring current begins to recover, the Dst begins a slow rise back to its quiet time level. The relationship of inverse proportionality between the horizontal component of the magnetic field and the energy content of the ring current is known as the Dessler-Parker-Sckopke relation. Other currents contribute to the Dst as well, most importantly the magnetopause current. The Dst index is corrected to remove the contribution of this current as well as that of the quiet-time ring current. [Hamilton *et.al.*, 1988]. The data was obtained from the World Data Center (WDC) for geomagnetism, Kyoto, Japan (<http://wdc.kugi.kyoto-u.ac.jp/dstae/>).

4.2.1.7 AE index

The Auroral Electrojet (AE) index was originally introduced by *Davis and Sugiura* [1966] as a measure of global electrojet activity in the auroral zone. The data was obtained from the WDC for geomagnetism, Kyoto, Japan (<http://wdc.kugi.kyoto-u.ac.jp/dstae/>). The AE index is derived from geomagnetic variations in the horizontal component observed at selected 12 observatories along the auroral zone in the northern hemisphere. To normalize the data, a base value for each station is first calculated for each month by averaging all the data from the station on the five international quietest days. This base value is subtracted from each value of one-minute data obtained at the station during that month.

4.2.1.8 Kp index

The daily mean of Kp index was obtained from the WDC for geomagnetism, Kyoto, Japan (<http://wdc.kugi.kyoto-u.ac.jp/kp/>). The Kp index was defined as the mean value of the disturbance levels in the two horizontal field components from 13 observatories located in the sub-auroral zone [Bartels, 2013]. There is a preliminary Kp index based on a subset of the observatories with about a 6 hours time lag. Because of the time resolution and the geographic positions of the observatories the Kp index can only serve as an overall measure of geomagnetic activity. The different storm phases during solar maximum (Gosling *et al.*, 1991) and during the declining phase of the solar cycle (Tsurutani *et al.*, 1995) are thus not resolved in detail using the Kp index.

4.2.1.9 Mesospheric temperature

The daily nighttime mean of mesospheric temperature data in Kelvin at height of 80-90 km was obtained from the Sounding of the Atmosphere with Broadband Emission Radiometry (SABER) on the NASA TIMED (Thermosphere-Ionosphere-Mesosphere Energetics and Dynamics) satellite data (<http://saber.gats-inc.com/>). The SABER is a 10 channel broadband, limb-viewing, infrared radiometer which has been measuring stratospheric and mesospheric temperatures since the launch of the TIMED satellite in December 2001. The TIMED satellite is in a nearly circular 625 km-altitude orbit with an inclination of 74°. The resulting latitudinal coverage extends over 135°, from 85° in one hemisphere to about 50° in the other. Every 60 days or so, the TIMED spacecraft executes a yaw maneuver which flips the dominant hemisphere covered by SABER. The temperature is obtained from the 15 mm radiation of CO₂. The SABER instrument provides measurements of the temperature at altitudes from ~16 to ~120 km with an accuracy of ~1–2K (Siskind *et al.*, 2005). This data is available in global observation but we only choose three paths namely NLK-CHF path (lat: 35°N-48°N, long: 139°E-121°W), NPM-CHF path (lat: 21°N-35°N, long: 139°E-159°W), and NWC-CHF path (lat: 21°S-35°N, long: 114°E-139°E).

4.2.1.10 Solar radio flux at 10.7 cm index

The daily mean of solar radio flux at 10.7 cm (2800 MHz) data in solar flux units (sfu, $1 \text{ sfu} = 10^{-22} \text{ W m}^{-2} \text{ Hz}^{-1}$) was obtained from the Interplanetary Magnetic Field (IMF) and plasma data (<http://omniweb.gsfc.nasa.gov/form/dx1.html>). The F10.7 index measurement is provided courtesy of the National Research Council Canada in partnership with the Natural Resources

Canada. The F10.7 radio emissions originate high in the chromosphere and low in the corona of the solar atmosphere. The F10.7 correlates well with the sunspot number as well as a number of Ultra Violet (UV) and visible solar irradiance records [King and Papitashvili, 2005].

4.2.1.11 Day of the year (DOY)

Seasonal variation can be learned using DOY from the first day to the end of each year. DOY is expressed as a combination of DOY1 and DOY2 in equation (4.1) and (4.2) [M. I. Nakamura et al., 2007].

$$DOY1 = \frac{\left(\sin\left(DOY \cdot \frac{2\pi}{365}\right) + 1\right)}{2} \quad (4.1)$$

$$DOY2 = \frac{\left(\cos\left(DOY \cdot \frac{2\pi}{365}\right) + 1\right)}{2} \quad (4.2)$$

where DOY is number from 1 to 365 (366).

4.2.1.12 Sunspot number (SSN)

Sunspot number is a quantity that measures the number of sunspots N_s and groups of sunspots N_g present on the surface of the sun on the Earth environment as defined in equation [Wolf, 1856].

$$R = k(10 \times N_g + N_s) \quad (4.3)$$

where k scaling coefficient, usually called the personal coefficient of the observer, allows compensating for the differences in the number of recorded sunspots by different ground observers. Sunspot number measurement is provided by solar influences data analysis center (SIDC), Royal Observatory of Belgium, Brussels Belgium. The data was obtained from <http://omniweb.gsfc.nasa.gov/form/dx1.html>.

4.2.2 Data Cleaning

The process of detecting and removing inaccurate data from the dataset is called data cleaning. The inconsistencies detected or removed may have been originally caused by user entry errors, by corruption in transmission or storage, or by different data dictionary definitions of similar entities in different stores. After removing or correcting the provided datasets the

results of cleansed time-series for each parameter then the dataset prepared for the sampling process.

4.2.3 Data Average

In this thesis, prediction of the nighttime mean of VLF electric field amplitude is a daily value. For this purpose, after the original dataset with incomplete or corrupt records are fixed with the linear interpolation from the original datasets, the new datasets are averaged into 2-min mean value and then take a mean value into daily nighttime mean values. Figure 4.4 is an example of VLF electric amplitude mid-latitude path (NPM-CHF path).

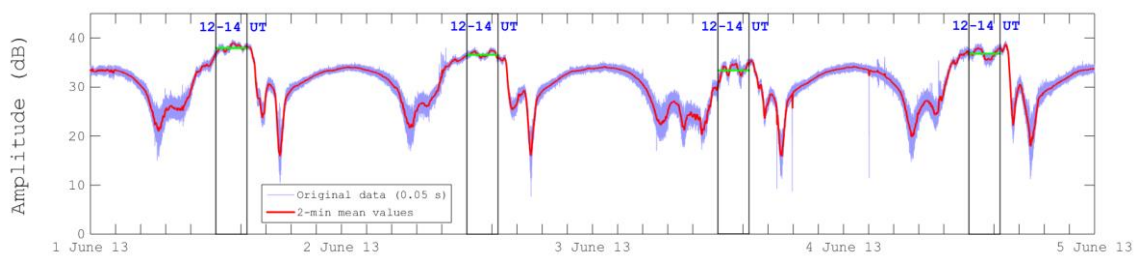


Figure 4.4: Time series of the VLF electric field amplitude from the VLF transmitter in Hawaii (NPM) received at Chofu Tokyo (CHF) for the time interval between 1 June 2013 and 4 June 2013. The clear diurnal pattern is recognized in the time series with a large amplitude during nighttime and a small amplitude in the daytime. The blue curve indicates the original VLF amplitude data with a temporal resolution of 0.05 s. The red curve indicates the 2-minute mean values of the original amplitude time series. The green horizontal line stands for the daily nighttime mean VLF amplitude by using the time interval of 12-14 UT (the local nighttime over the NPM-CHF path) represented by the black box.

Figure 4.4 demonstrates the example of the VLF electric field amplitude time series with the time interval from 1 June 2013 to 4 June 2013 (4 days). The original time series of VLF electric field amplitude with a time resolution of 0.05 s (sampling rate of 20 Hz) at receiver station in Chofu, Tokyo, Japan correspond to NPM-CHF path shown in blue curve. Corresponding 2-minute mean value of the original VLF data (down-sampled data) is shown in red curve. The green horizontal line indicates the daily nighttime mean VLF amplitude calculated by taking mean value from 2-minute value between the time interval from 12-14 UT represented by the black box. As is seen from the figure 3.3, electric field amplitude has a clear and typical diurnal pattern of subionospheric VLF wave propagation with several sharp drops indicating terminator times due to mode interferences. Time interval with smooth and smaller amplitude value is for the local daytime period (about from -5 UT to +5 UT) due to a large absorption of the signal, whilst the time interval with more variability (than daytime) with

larger amplitude value is for the local nighttime period. The horizontal lines clearly represent and track the daily nighttime amplitude. We use daily nighttime value (one point per day) for data analysis but we connect the daily value by linear lines for readers to track easily the day-to-day variability in Figure 4.4.

4.2.4 Data Normalization

The process of simplifying the design of a dataset so that it achieves the optimal structure and also increasing the efficiency of the network is called data normalization. For this purpose, all the dataset normalized and the feature scaling method is adopted and transferred all the dataset values with the range from -1 to 1 as shown in equation (4.4)

$$x_N = 2 \cdot \frac{x - x_{min}}{x_{max} - x_{min}} - 1 \quad (4.4)$$

where x is a original value of dataset before normalization applied as the initial value, x_N represents the normalized value, and x_{min} and x_{max} represent the minimum and maximum value of the initial dataset. The purpose of the normalization step is integration process may produce very high results if the original datasets are fed into the neuron of the network and caused the tangent sigmoid transfer function to perform in the poor performance to determine small changes in the inputs data also lack of sensitivity.

4.3 Training Strategy

Neural Network (NN) is a computational paradigm inspired from the structure of biological NN and their way of encoding and solving problems [Ratliff, 1968; Rumelhart *et al.*, 1988; Tulunay *et al.*, 2004]. A NN is able to identify underlying highly complex relationships based on input-output data only. In this thesis, we use three NN training algorithms: a Levenberg Marquardt Neural Network (LMANN) [Hagan and Menhaj, 1994], a Bayesian Neural Network (BRANN) [Foresee and Hagan, 1997] and a Scaled Conjugate Gradient (SCG) [Møller, 1993]. The architecture of the network must be decided first for developing the required NN to model the data. The NN consists of a combination of a number of hidden and output layers. The hidden layers perform the mapping between inputs and outputs to the network in a feedforward arrangement.

A NARX NN is a dynamic NN and contains recurrent feedbacks from several layers of the network to the input layer with time-delayed units. In addition, a NARX NN is trained to

capture the relationship between the predicted values, residuals and original time series. The weights and biases of NARX NN are kept to predict the future value of original time series [Ardalani and Zolfaghari, 2010]. The long-term VLF daily nighttime electric field amplitude was trained by using a series-parallel approach and the NARX model represented by a discrete-time nonlinear system is shown by equation (4.5) [Chen *et al.*, 1989].

$$y(k+1) = F[(y(k), y(k-1), \dots, y(k-d_y+1); u(k), u(k-1), \dots, u(k-d_u+1))] \quad (4.5)$$

where $u(k) \in \mathbb{R}$ and $y(n) \in \mathbb{R}$ denote, respectively, the input and output of the model at time step k , while $d_u \geq 1$ and $d_y \geq 1, d_u \geq d_y$, are the input-memory and output-memory orders, respectively.

The data structure of both u and y for the NARX model are in the form of a continuous time-series sequence as shown in equation (4.6)

$$\begin{cases} u = \{[u_1] [u_2] \dots [u_k]\} \\ y = \{[y_1] [y_2] \dots [y_k]\} \end{cases} \quad (4.6)$$

where the element of u and y , i.e. $[u_t]$ and $[y_t]$, respectively, are the values collected at a given time point t ($1 \ll t \ll k$). By configuring the TDL in advance such as fixing the value of $d_y + 1$ and $d_u + 1$ in Eq. (1) with the input data, the NARX model can be establish for time-serials prediction. F is nonlinear tangent sigmoid function given by equation (3.2) [Menon *et. al.*, 1996].

$$F(x) = \frac{1}{1 + \exp(-x)} - \frac{1}{2} \quad (4.7)$$

The mathematical representation of NARX NN structure can be produced by combining using recurrent NN architecture is given by equation (4.7) [Ugalde *et. al.*, 2014].

$$y(k) = X^* \varphi_2 \{ (V_B \varphi_1 (J_u W_B) + V_A \varphi_1 (J_y W_A) + b \varphi_1) + Z_H \} \quad (4.8)$$

where:

$$J_u = [u(k)u(k-1) \dots u(k-d_u+1)] \in R^{1 \times d_u}$$

$$J_y = [y(k)y(k-1) \dots y(k-d_y+1)] \in R^{1 \times d_y}$$

d_u is the number of pass inputs of the system and d_y is the number of pass outputs of the system

$$W_B = [W_{b_{i,1}} \ W_{b_{i,2}} \ \dots \ W_{b_{i,d_u+1}}]^T \in R^{1 \times d_u}$$

$$W_A = [W_{a_{i,1}} \ W_{a_{i,2}} \ \dots \ W_{a_{i,d_y+1}}]^T \in R^{1 \times d_y}$$

$$X, V_B, V_A, Z_H \in R^1$$

$X, V_B, V_A, b, Z_H, W_{b_i}$ and W_{a_i} are the synaptic weights. J_u and J_y are input and output regressor vectors. φ_1 and φ_2 are the activation functions (linear or nonlinear) of the NN. $i = 1, 2, \dots, nn$ and nn is the number of neurons. d_u is the number of pass inputs of the system and d_y is the number of pass outputs of the system. For evaluating the performance of a network, performance criteria are chosen to observe the error between the desired responses (original data) and the calculated outputs (prediction). The RMSE is the root mean squared of the error between original data and predicted value. When the RMSE is used as the performance criterion, it is implicitly assumed that the errors have Gaussian distribution. The minimization of the RMSE is to obtain the best performance of the network. Definition of RMSE is given by equation (4.9)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (error_{sim}(i))^2} \quad (4.9)$$

where n is the number of data points (one point per day), and $error_{sim}(i)$ is the error difference (in dB) between the output datasets from observation and the model prediction at the certain data point (day i).

4.3.1 Time Delay Selection

Figure 4.5: show time delay selection effect in the different algorithms with the example of 200 neurons in the hidden layer and increasing time delay from 3 days to 10 days before the given day related with the performance of Pearson correlation coefficient. The BRANN algorithm is inferior compared with the LMANN algorithm. Pearson correlation coefficient (r) for the BRANN algorithm is smaller than the LMANN algorithm. Pearson

correlation coefficient increases with increasing the delay time but not as high as the LMANN algorithm. In contrast, The SCG has the worst performance among the other algorithms. It is indicated by the smallest correlation coefficient. The LMANN is the best algorithm with the highest correlation and continues increases with the delay time increases.

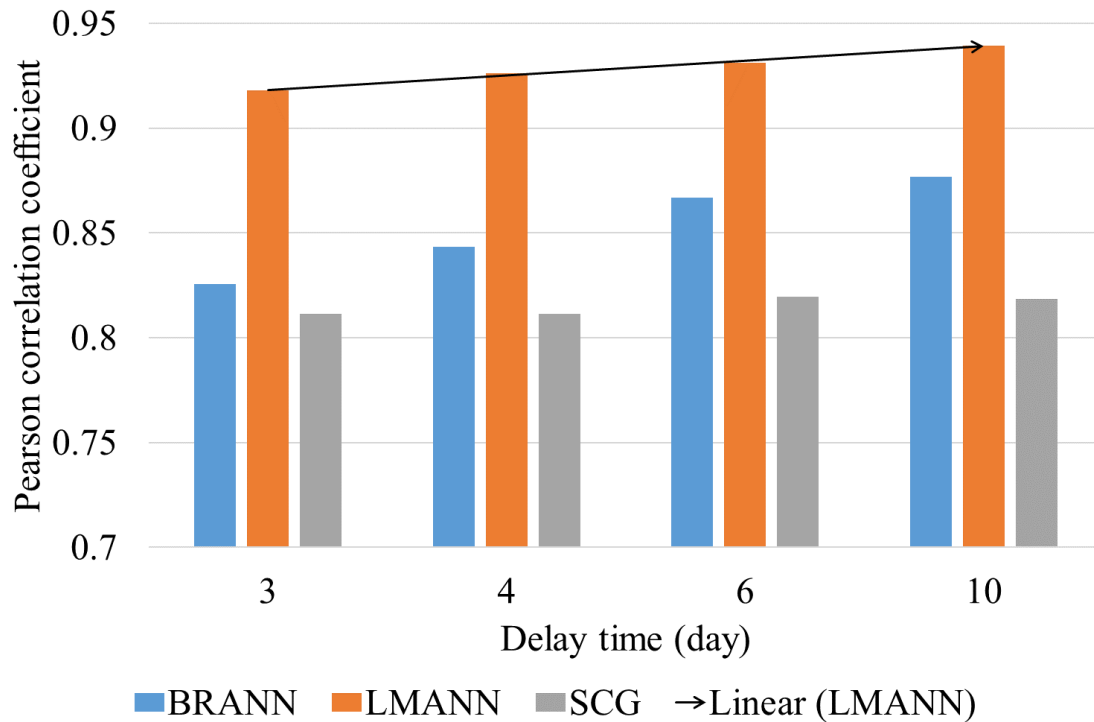


Figure 4.5: Performance of three different training algorithms based on Pearson correlation coefficient with time delay selection.

Figure 4.6 shows time delay selection effect in the different algorithms with the example of 200 neurons in the hidden layer and increasing time delay from 3 days to 10 days before the given day associated with RMSE. The RMSE for the BRANN algorithm is higher than the LMANN algorithm. RMSE slightly decreases with increasing the delay time but not as much as the LMANN algorithm. In contrast, The SCG has the worst performance among the other algorithms. It is indicated by highest RMSE. And still, the LMANN is the best algorithm with the smallest error and continue increases with the delay time increases.

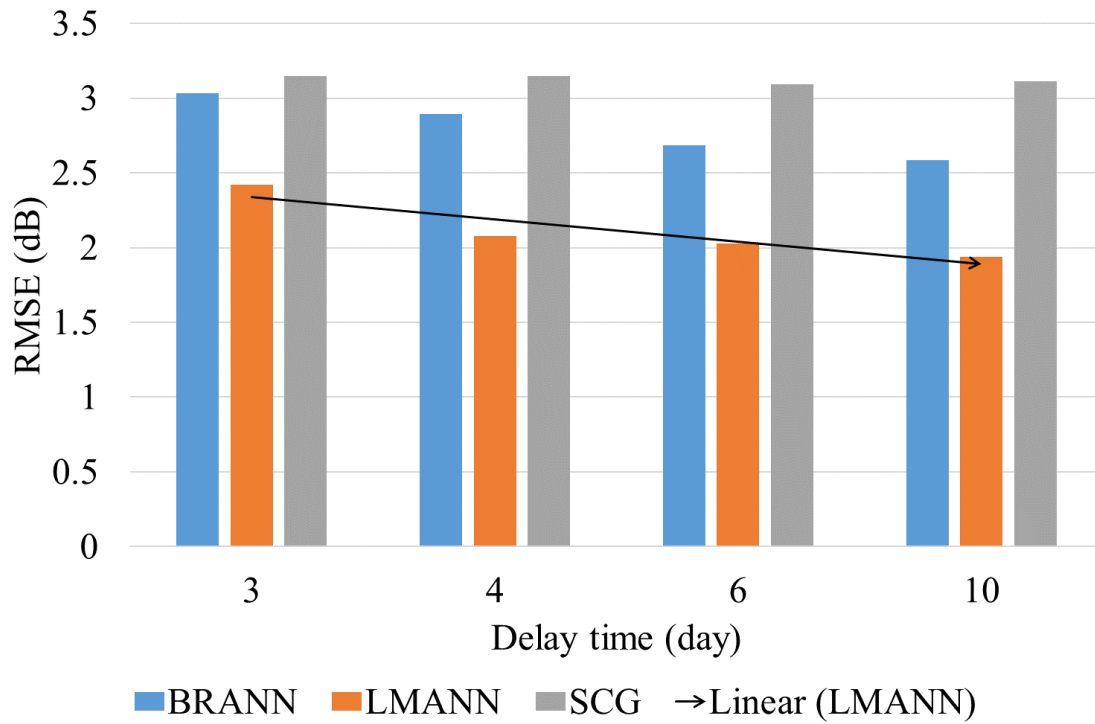


Figure 4.6: Performance of three different training algorithms based on RMSE with time delay selection.

4.3.2 Hidden Layer Size

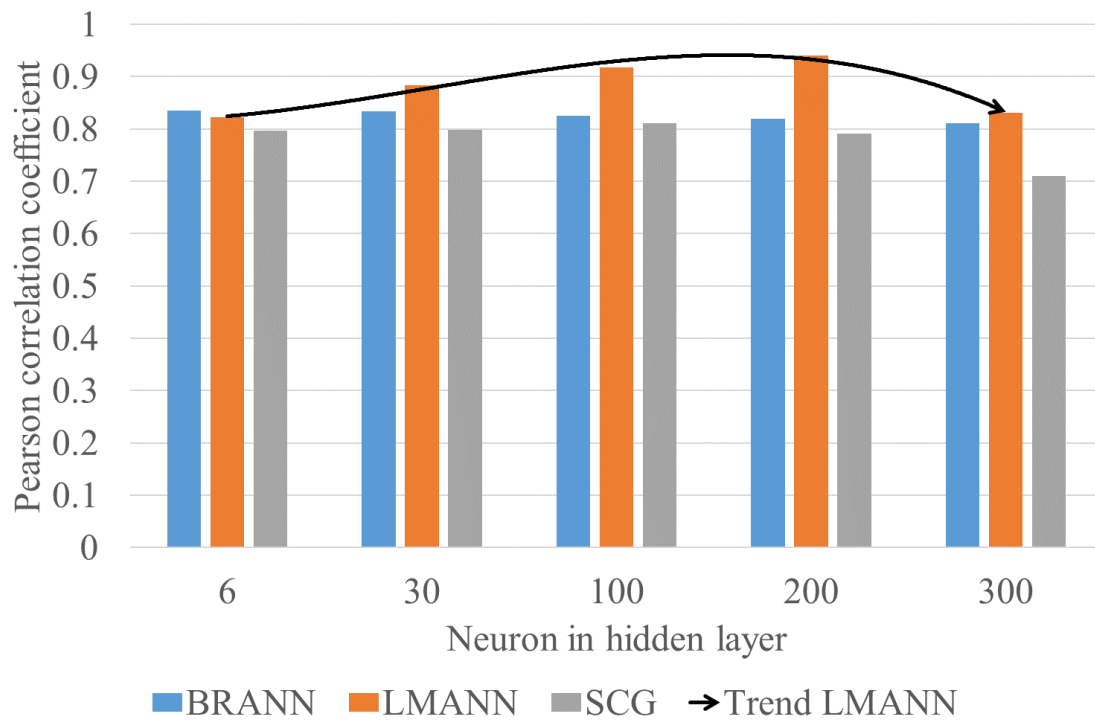


Figure 4.7: Performance of three different training algorithms based on Pearson correlation coefficient with hidden layer size selection.

Figure 4.7 show hidden layer size selection effect in the different algorithms with the example of 3 days delay time and increasing number of neurons in the hidden layer from 6 to 300 neurons related with the performance of Pearson correlation coefficient. The BRANN is relatively stable with the average of Pearson correlation coefficient (r) around 0.8. Pearson correlation coefficient slightly decreases with increasing the number of neurons in the hidden layer but not as high as the LMANN algorithm. In contrast, The SCG has performance quite similar with the BRANN. The LMANN is the best algorithm with the highest correlation and continues increases with increasing neuron number in the hidden layer. But after used 300 neurons, the performance decreasing.

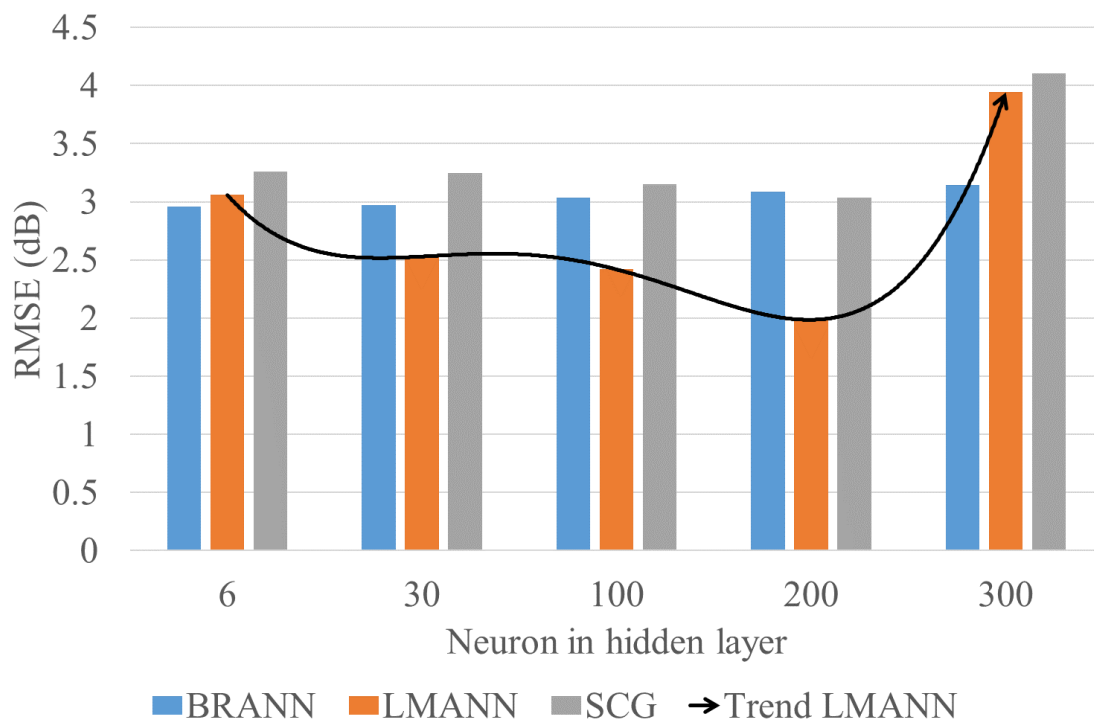


Figure 4.8: Performance of three different training algorithms based on RMSE with hidden layer size selection.

Figure 4.8 show hidden layer size selection effect in the different algorithms with the example of 3 days delay time and increasing number of neurons in the hidden layer from 6 to 300 neurons related with the performance of RMSE. The error prediction of BRANN algorithm is relatively stable with the average of RMSE around 3 dB. The SCG has performance quite similar with the BRANN, but in the neurons number of 300 the RMSE increasing significantly. The LMANN is the best algorithm with the smallest RMSE and continues decreases with increasing neuron number in the hidden layer even though in 300 neurons also increasing like SCG algorithm.

4.3.3 Training Algorithm

Figure 4.9 summarizes the Pearson correlation coefficient (r) of the three different training algorithms used in this thesis with different parameter settings. The LMANN algorithm has the best performance among the three algorithms with the largest Pearson correlation coefficient (r) with the tendency increases. The correlation coefficient for the LMANN with the three days of input-memory increases by the increasing the number of neurons in the hidden layer from six to two hundred neurons. However, the correlation coefficient then decreases when increasing the number of neurons in the hidden layer above two hundred.

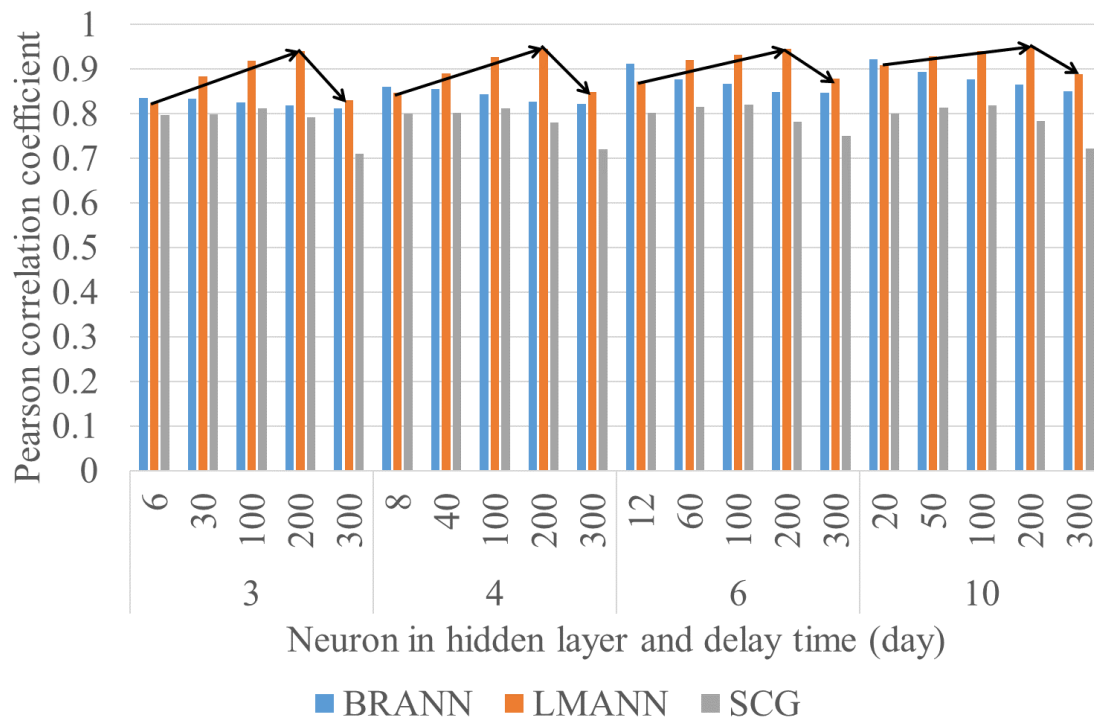


Figure 4.9: The best algorithm based on Pearson correlation coefficient with the variables changing in neurons number in the hidden layer and time delay.

Figure 4.10 The LMANN algorithm also has the best performance among the three algorithms with the smallest RMSE. The RMSE tends to decrease with increasing neuron numbers and reaches a minimum value at two hundred, before increasing again at three hundred neurons in the hidden layer.

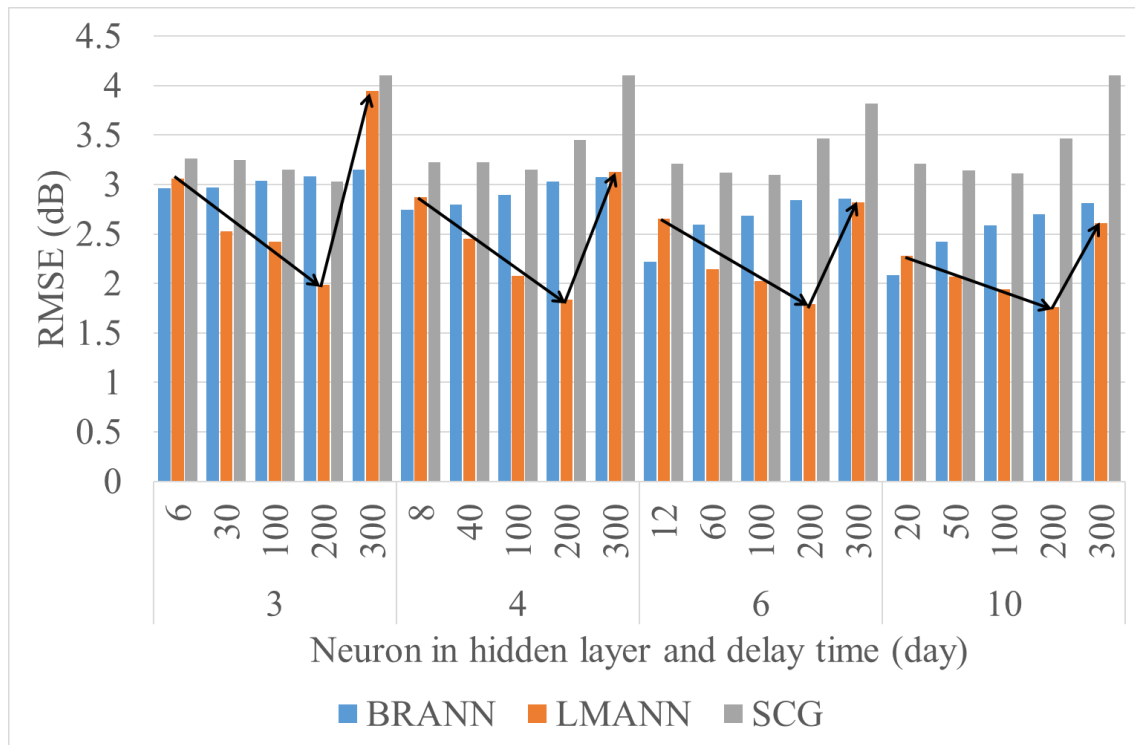


Figure 4.10: The best algorithm based on RMSE with the variables changing in neurons number in the hidden layer and time delay.

4.4 The Best Network Selection

Table 4.1 summarizes the performance of the three different training algorithms used in this paper with different parameter settings. The LMANN algorithm has the best performance among the three algorithms with the largest Pearson correlation coefficient (r) and the smallest RMSE. The RMSE tends to decrease with increasing neuron numbers and reaches a minimum value at 200, before increasing again. For example, the correlation coefficient for the LMANN with the 3 days of input memory increases from 0.8226 to 0.9407 by increasing the number of neurons in the hidden layer from 6 to 200 neurons. However, the correlation coefficient then decreases when increasing the number of neurons in the hidden layer above 200. The correlation coefficient decreases to 0.8306 for the case of 300 neurons in the hidden layer. Thus, LMANN algorithm shows the best performance with the largest correlation coefficient r (the bolded values in Table 1) at 200 neurons in the hidden layer even with different days in the input memory, which has a variability $\Delta r \sim 0.008$ (0.9407–0.9488). The corresponding variability of RMSE between the smallest and highest input memory of the system is small as shown by $\Delta \text{RMSE} \sim 0.22 \text{ dB}$ (from 1.9853 to 1.7574 dB).

On the other hand, the performance of the BRANN algorithm is inferior with that of LMANN algorithm. Pearson correlation coefficient (r) for the BRANN algorithm is smaller, and the RMSE is greater than those for the LMANN algorithm. Pearson correlation coefficient decreases with increasing the number of neurons in the hidden layer. For example, the BRANN algorithm with a 10-day input memory and 20 neurons in the hidden layer has a Pearson correlation coefficient (r) of 0.9219, but the correlation decreases with increasing neuron number, such that 300 neurons lead to an r of 0.8504. Also, r increases with increasing the number of input memory. However, RMSE significantly decreases with increases in the number of days in the input memory from 3.0855 dB to 2.7019dB for 200 neurons in the hidden layer.

Table 4.1 Fit Pearson correlation coefficient (r) and performance error (RMSE) of each model.

Input-memory (days)	Neurons in The Hidden Layer	Pearson Correlation Coefficient (r)			Root Mean Square Error (RMSE)		
		BRANN	LMANN	SCG	BRANN	LMANN	SCG
3	6	0.8349	0.8226	0.7965	2.9605	3.0588	3.2604
	30	0.8340	0.8835	0.7977	2.9688	2.5290	3.2435
	100	0.8258	0.9181	0.8116	3.0333	2.4188	3.1483
	200	0.8191	0.9407	0.7914	3.0855	1.9853	3.0330
	300	0.8111	0.8306	0.7101	3.1462	3.9427	4.1040
4	8	0.8605	0.8461	0.8000	2.7415	2.8713	3.2267
	40	0.8542	0.8904	0.8010	2.7976	2.4506	3.2235
	100	0.8433	0.9261	0.8115	2.8918	2.0794	3.1496
	200	0.8262	0.9449	0.7797	3.0310	1.8386	3.4510
	300	0.8209	0.8476	0.7201	3.0724	3.1242	4.1022
6	12	0.9112	0.8718	0.8014	2.2153	2.6536	3.2124
	60	0.8765	0.9205	0.8147	2.5912	2.1472	3.1166
	100	0.8669	0.9313	0.8196	2.6831	2.0260	3.0933
	200	0.8488	0.9459	0.7817	2.8439	1.7878	3.4685
	300	0.8468	0.8784	0.7505	2.8599	2.8174	3.8150
10	20	0.9219	0.9081	0.8004	2.0837	2.2758	3.2105
	50	0.8928	0.9280	0.8126	2.4192	2.0665	3.1389
	100	0.8769	0.9394	0.8186	2.5838	1.9399	3.1144
	200	0.8642	0.9488	0.7825	2.7019	1.7574	3.4631
	300	0.8504	0.8883	0.7220	2.8126	2.6074	4.1041

The SCG has the worst performance among the three algorithms considered in this paper. It is indicated by smallest correlation coefficient and the largest RMSE in comparison to the BRANN and LMANN algorithms. For example, the SCG with the 3 days of input memory, the correlation increases from 0.7965 to 0.8116 by increasing the number of neurons in the hidden layer from 6 to 100 neurons. However, the correlation coefficient decreases by increasing the number of neurons in the hidden layer greater than 100. For SCG with the 3 days

of input memory, the correlation coefficient decreases from 0.8116 to 0.7101 with increasing the number of neurons from 100 to 300 neurons in the hidden layers, respectively. The RMSE does not significantly change with increasing input memory of the system. RMSE with 3 days of input memory slightly decreases from 3.1483 dB to 3.1144dB with increasing input memory to 10 days for 100 neurons in the hidden layer.

Among the three algorithms, we consider that the LMANN algorithm is found to have the best performance, with the largest correlation and the lowest prediction error for each input memory in the system as shown in Table 4.1 with the bolded values. Moreover, the performance of the LMANN algorithm does not vary significantly with increasing number of neurons in the hidden layer and input memory, which enable us to maintain the high prediction performance with a relatively small number of neurons in the hidden layer and input memory that allows minimizing the computational cost.

4.5 Conclusions

This chapter deals with the general methodology for dynamic modeling of VLF amplitude and critical frequency in F2-region (foF2). In this case, this approach is focused on modeling the nonlinear dynamic changing of the daily nighttime mean value of VLF electric amplitude in mid-latitude path received by Chofu station and foF2 data from Kokubunji station. This methodology is based on the knowledge about the VLF wave propagation and ionosonde.

Further, the first nonlinear models relate the characteristic of VLF radio wave propagation in the Earth-ionosphere cavity, external forcings from below and above the ionosphere. Second, characteristic of F-region to high frequency (HF) radio propagation and also external sources influencing in this region. Thus, the main conclusion extracted from this relationship is that the VLF amplitude changing in the mid-latitude path is related with stratospheric temperature, cosmic rays, total column ozone, Dst, AE and Kp-indices, mesospheric temperatures, and F10.7 index. Moreover, for foF2 is related to DOY1, DOY2, F10.7, Dst, AE, and Kp indices, SSN.

Concerning the selection of the NARXNN methodology such as training algorithm, time delay, and hidden layer size are decided based on the results of the training process. Thus, based on the results among the three algorithms (BRANN, LMANN, and SCG) we consider that the LMANN algorithm with delay time three days before the given day and two hundred neurons in the hidden layer is found to have the best performance, with the largest correlation and the lowest prediction error. Moreover, the performance of the LMANN algorithm does not

vary significantly with increasing number of neurons in the hidden layer and input memory, which enable us to maintain the high prediction performance with a relatively small number of neurons in the hidden layer and input memory that allows minimizing the computational cost.

The constructed NARXNN model indicates that stratospheric temperature two days before the given day has the most significant contribution to VLF amplitude variation, as well as the VLF amplitude from the previous day. Total column ozone and Dst index also significantly influence the VLF amplitude from two days before and previous day from given day, respectively.

CHAPTER 5

5. Applications for VLF Electric Field Amplitude Time Series Data

The results from Chapter 5 that the NARXNN model for VLF electric field amplitude mid-latitude path (NPM-CHF) using time delay of three days before the given days and three hundred neurons in the hidden layer with LMANN training algorithm offers the powerful nonlinear modeling and predicting indicate the highest Pearson correlation coefficient and smallest RMSE. It was shown in Chapter 4 that recurrent neural network is capable of learning short-time dependencies due to their global feedback from the output of the network to the network input. In contrast, NARXNN is capable of learning long-term dependencies. This implies that the output at the present time is dependent on the present and past values of the input as well as the past values of the output itself. *Bengio et al.* [1994] show the reason for learning long-term dependencies is not a feasible task.

Furthermore, NARXNN can be used for two objectives: one-step ahead prediction and multi-step ahead prediction tasks. The one step prediction of NARXNN model provides highly accurate predictions and suitable for the nonlinear dynamic system. Furthermore, the multi-step ahead prediction task becomes a dynamic modeling task, in which the neural network model acts as an autonomous system, trying to recursively emulate the dynamic behavior of the system that has generated the nonlinear time series. Moreover, the dynamic modeling and multi-step ahead prediction are much more complex to deal with than one-step ahead prediction, and in particular, recurrent neural architectures, these are complex tasks in which neural network models play an important role [*Hu and Wang, 2006*]. So in the multi-step ahead prediction, a prediction iterated for k times returns a k -step ahead forecasting.

In this chapter, we developed an operational tool for predicting daily nighttime mean amplitude of the subionospheric VLF propagation from one day ahead up to ten days ahead at Chofu (CHF), Tokyo, Japan (35.6°N, 139.5°E) and Tsuyama (TYM), Japan (35.1°N, 133.9°E) by using NARXNN. In this regard, there could be several indices describing VLF amplitude in the nighttime such as stratospheric and mesospheric temperatures, cosmic rays, total column ozone, F10.7, Dst, AE and Kp-indices. To find the significantly important parameters, we analyzed past VLF nighttime amplitude in a period of three years for CHF station and around

two years for TYM station in three different latitude paths namely low-, mid-, and high-latitude paths by using NARXNN.

5.1 Inputs/output Data Processing

The selection of the input datasets was influenced by the physical studies carried out to understand the physical meaning of the relationship between VLF electric field amplitude changing and lower ionospheric condition in the different latitude paths. The data inputs datasets into the NARXNN were daily nighttime mean values of VLF electric field amplitude stratospheric temperature, cosmic rays, total column ozone, Dst, AE and Kp-indices, mesospheric temperatures, and F10.7 index in three different latitude paths associated with CHF and TYM receiving stations. Before the datasets used in the training and prediction, there was preprocessing data as explain in Section 4.2.

Table 5.1 shows the interval time of input dataset for VLF amplitude prediction model both at Chofu and Tsuyama VLF receiver stations. The input dataset for Chofu station with the time interval between 1 January 2011 and 31 December 2013 (1096 data points). The training period used a data from 1 January 2011 to 4 February 2013 (766 data points) and the prediction period between 5 February 2013 to 31 December 2013 (330 data points). Tsuyama station has a different dataset with time interval from 15 March 2014 to 26 May 2016 (804 data points). Further, divided into two periods, between 15 March 2014 and 28 September 2015 (563 data points) is called training period and prediction period from 29 September 2015 to 26 May 2016 (241 data points).

Table 5.1: The interval data of daily VLF prediction model for training and prediction periods

Description	Training period	Prediction period
Chofu station	01-01-2011 to 04-02-2013 (766 data points)	05-02-2013 to 31-12-2013 (330 data points)
Tsuyama station	15-03-2014 to 28-09-2015 (563 data points)	29-09-2015 to 26-05-2016 (241 data points)

5.1.1 Low-Mid-Latitude VLF Propagation Path

Figure 5.1 illustrates daily nighttime mean values of the target and inputs of NARXNN for the low-mid-latitude (NWC-CHF) path over the time interval from 1 January 2011 to 31 December 2013. This figure is derived from the raw data with range time between

15 UT and 19 UT. Figure 5.1(a) shows the nighttime VLF amplitude received at CHF over three years from NWC VLF transmitter station at Harold E. Holt, North West Cape, Exmouth, Australia. This variation appears to have some local maxima and minima for summer and winter times respectively in the northern hemisphere, however it indicates a complex response. Figure 5.1(b) shows the daily mean of the nighttime stratospheric temperature time series obtained from the AIRS instrument measuring during the orbits of descending node. Figure 5.1(c) shows the daily intensity of the cosmic ray counts as time series. The daily mean of total column ozone (TCO) is shown Figure 5.1(d). Furthermore, Figure 5.1(e), Figure 5.1(f) and Figure 5.1(g) show the time series of Dst, AE and Kp indices respectively. The daily mean mesospheric temperature is shown Figure 5.1(h). Finally, Figure 5.1(i) shows the time series of F10.7 index. As seen from Figure 5.1, the temporal dependences of the TCO, stratospheric and mesospheric temperatures appear to have similar seasonal dependences but with some timing offset, while the cosmic ray has a systematic descending trend over three years presumably due to modulation from changing solar activity.

Figure 5.2 shows daily nighttime mean values of the target and inputs of NARXNN for the low-mid-latitude (NWC-TYM) path over the time interval from 15 March 2014 to 26 May 2016. This figure is derived from the raw data with range time between 15 UT and 19 UT and the hour-time interval same as Chofu station for low-mid-latitude path. The nighttime VLF amplitude received at Tsuyama VLF receiver station over three years from NWC VLF transmitter station at Harold E. Holt, North West Cape, Exmouth, Australia as shown in Figure 5.2(a). Then, Figure 5.2(b) illustrates the daily mean of the nighttime stratospheric temperature time series at 10 hPa (~ 30 km). Figure 5.2(c) depicts the daily intensity of the cosmic ray counts as time series at Magadan station. The daily mean of total column ozone (TCO) from the Ozone Monitoring Instrument (OMI) is shown Figure 5.2(d). Furthermore, Figure 5.2(e), Figure 5.2(f) and Figure 5.2(g) show the time series of Dst, AE and Kp indices respectively were obtained from world data center in Kyoto station. The daily mean mesospheric temperature in Kelvin at altitude of 80 – 90 km is shown Figure 5.2(h). Finally, Figure 5.2(i) shows the time series of F10.7 index. As seen from Figure 5.2, the temporal dependences of the TCO and stratospheric temperature appear to have similar seasonal dependences but with some timing offset.

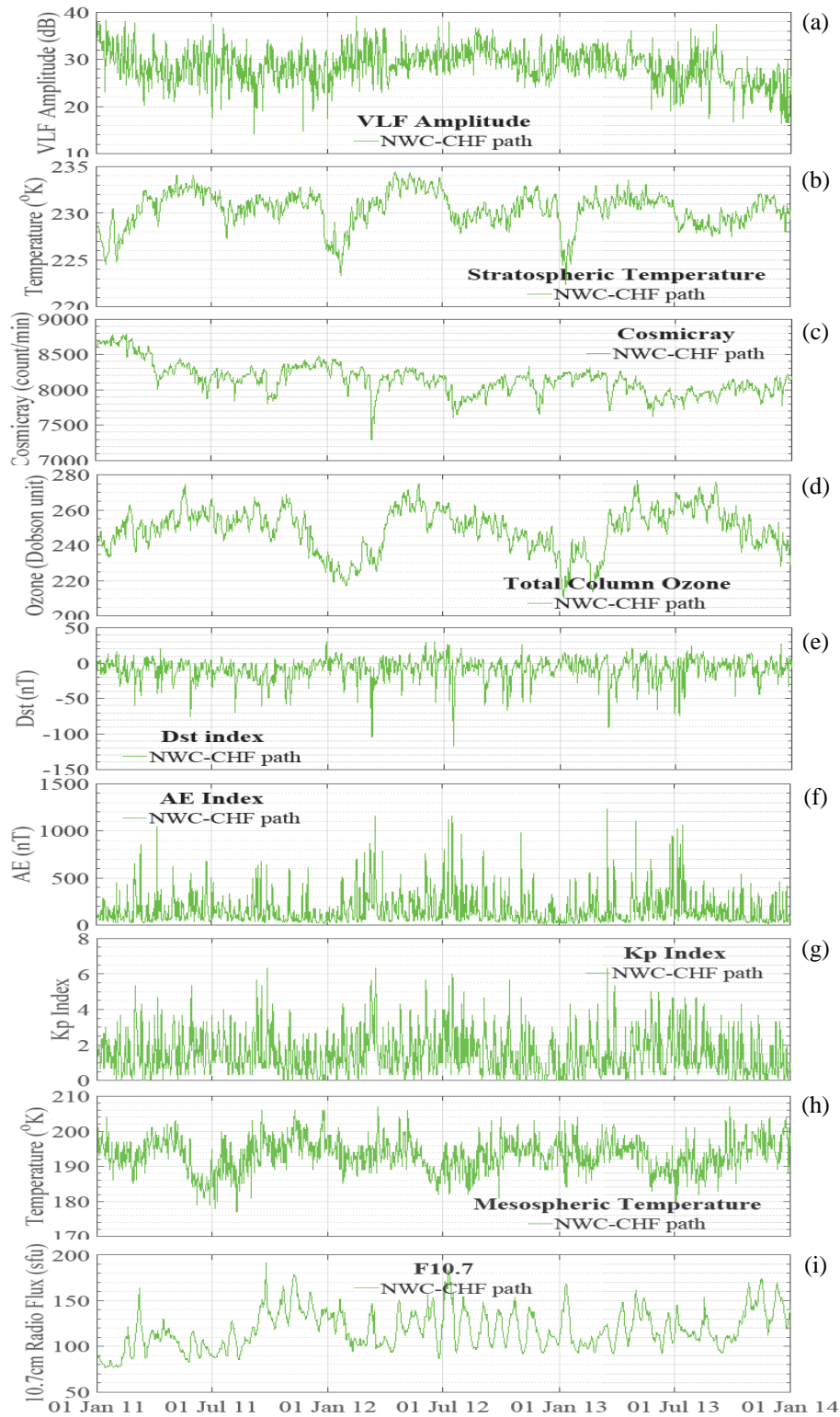


Figure 5.1: Daily nighttime mean values of the target and inputs of NARXNN for low-mid-latitude path from Cape town, Australia (NWC) to Chofu, Japan (CHF) station over the time interval from 1 January 2011 to 31 December 2013. (a) VLF daily nighttime mean amplitude for NWC-CHF path as an input and target of the system (dB), and seven parameters as inputs to the system, (b) stratospheric temperature ($^{\circ}$ K), (c) cosmic ray (in count/ min), (d) total column ozone (Dobson unit), (e) Dst index (nT), (f) AE index (nT), (g) Kp index, (h) mesospheric temperature ($^{\circ}$ K), and (i) F10.7 index (sfu).

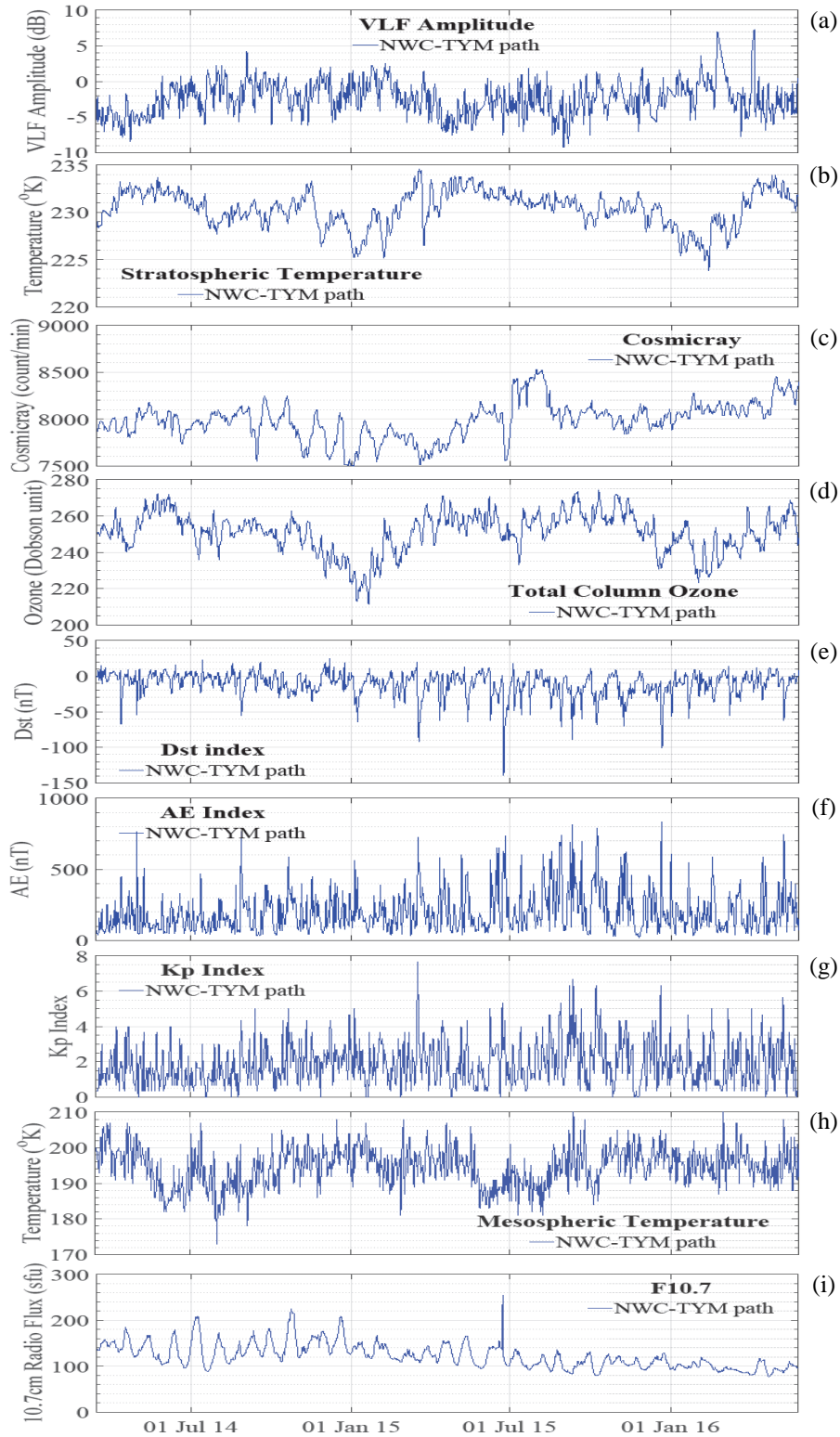


Figure 5.2: Daily nighttime mean values of the target and inputs of NARXNN for low-mid-latitude path from Cape town, Australia (NWC) to Tsuyama, Japan (TYM) station over the time interval from 15 March 2014 to 26 May 2016. (a) VLF daily nighttime mean amplitude for NWC-TYM path as an input and target of the system (dB), and seven parameters as inputs to the system, (b) stratospheric temperature ($^{\circ}\text{K}$), (c) cosmic ray (in count/ min), (d) total column ozone (Dobson unit), (e) Dst index (nT), (f) AE index (nT), (g) Kp index, (h) mesospheric temperature ($^{\circ}\text{K}$), and (i) F10.7 index (sfu).

5.1.2 Mid-Latitude VLF Propagation Path

Figure 5.3 illustrates daily nighttime mean values of the target and inputs of NARXNN for the mid-latitude (NPM-CHF) path over the time interval from 1 January 2011 to 31 December 2013. This figure is derived from the raw data with range time between 12 UT and 14 UT. Figure 5.3(a) shows the nighttime VLF amplitude received at CHF over three years from NPM VLF transmitter station at Pearl Harbour, Lualuahei, Hawaii, United States of America. This variation appears to have some local maxima and minima for summer and winter times respectively in the northern hemisphere, however it indicates a complex response. Figure 5.3(b) shows the daily mean of the nighttime stratospheric temperature time series in Kelvin at a height of ~ 30 km obtained from the AIRS instrument measuring during the orbits of descending node. The daily intensity of the cosmic ray counts per minute obtained from Magadan cosmic ray station with geographical location of 60.04°N , 151.05°E shown in Figure 5.3(c). The daily mean of total column ozone (TCO) obtained from the Ozone Monitoring Instrument (OMI) aboard on the NASAs AURA spacecraft is shown Figure 5.3(d). Furthermore, Figure 5.3(e), Figure 5.3(f) and Figure 5.3(g) show the time series of Dst, AE and Kp indices respectively were was obtained from the world data center for geomagnetism, Kyoto, Japan. The daily mean mesospheric temperature in Kelvin at height of 80-90 km was obtained from the Sounding of the Atmosphere with Broadband Emission Radiometry (SABER) on the NASA TIMED (Thermosphere-Ionosphere-Mesosphere Energetics and Dynamics) satellite is shown Figure 5.3(h). Finally, Figure 5.3(i) shows the time series of F10.7 index and provided courtesy of the National Research Council Canada in partnership with the Natural Resources Canada. As seen from Figure 5.3, the temporal dependences of the TCO, stratospheric and mesospheric temperatures appear to have similar seasonal dependences but with some timing offset, while the cosmic ray has a systematic descending trend over three years presumably due to modulation from changing solar activity.

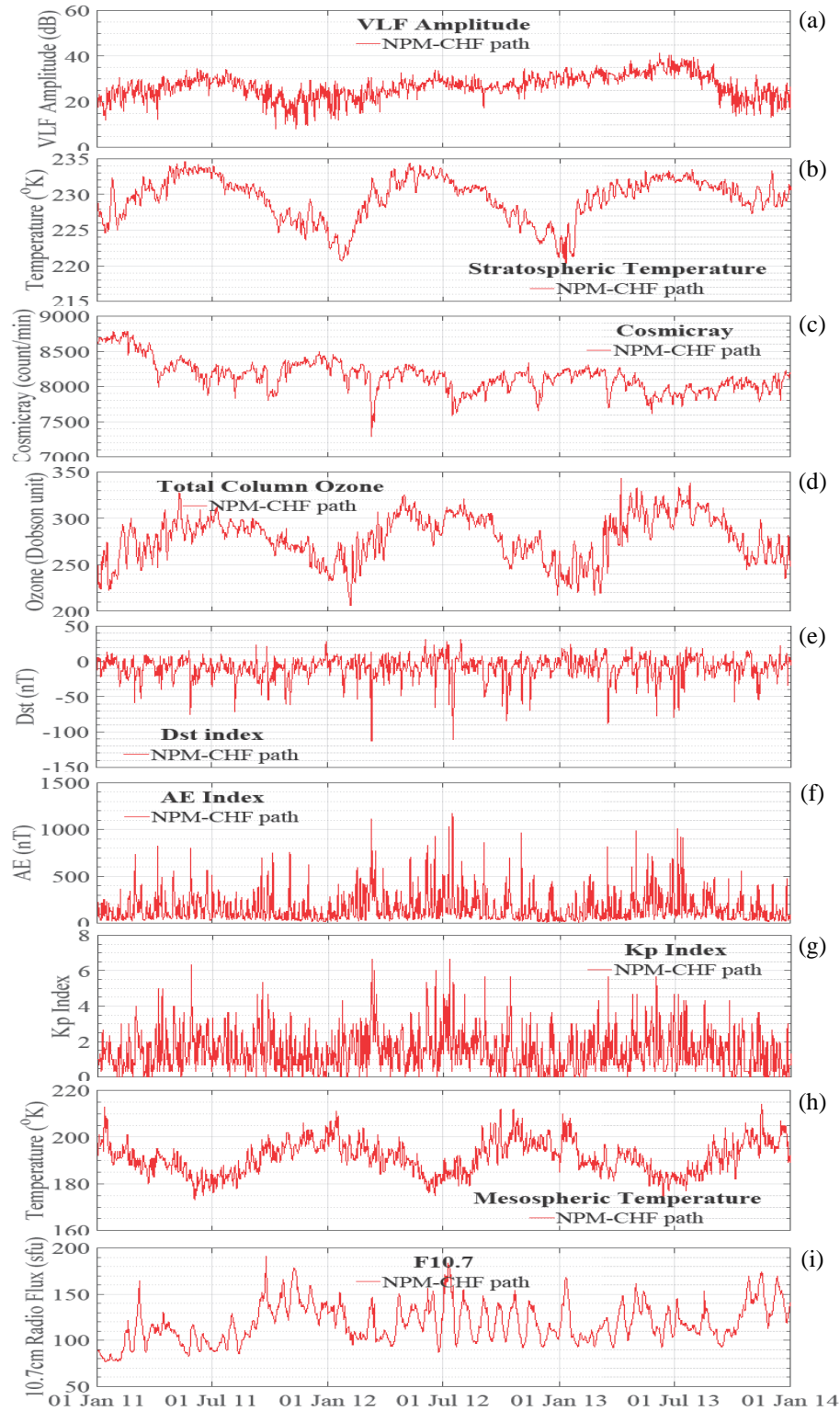


Figure 5.3: Daily nighttime mean values of the target and inputs of NARXNN for mid-latitude path from Hawaii, USA (NPM) to Chofu, Japan (CHF) station over the time interval from 1 January 2011 to 31 December 2013. (a) VLF daily nighttime mean amplitude for NPM-CHF path as an input and target of the system (dB), and seven parameters as inputs to the system, (b) stratospheric temperature ($^{\circ}\text{K}$), (c) cosmic ray (in count/ min), (d) total column ozone (Dobson unit), (e) Dst index (nT), (f) AE index (nT), (g) Kp index, (h) mesospheric temperature ($^{\circ}\text{K}$), and (i) F10.7 index (sfu).

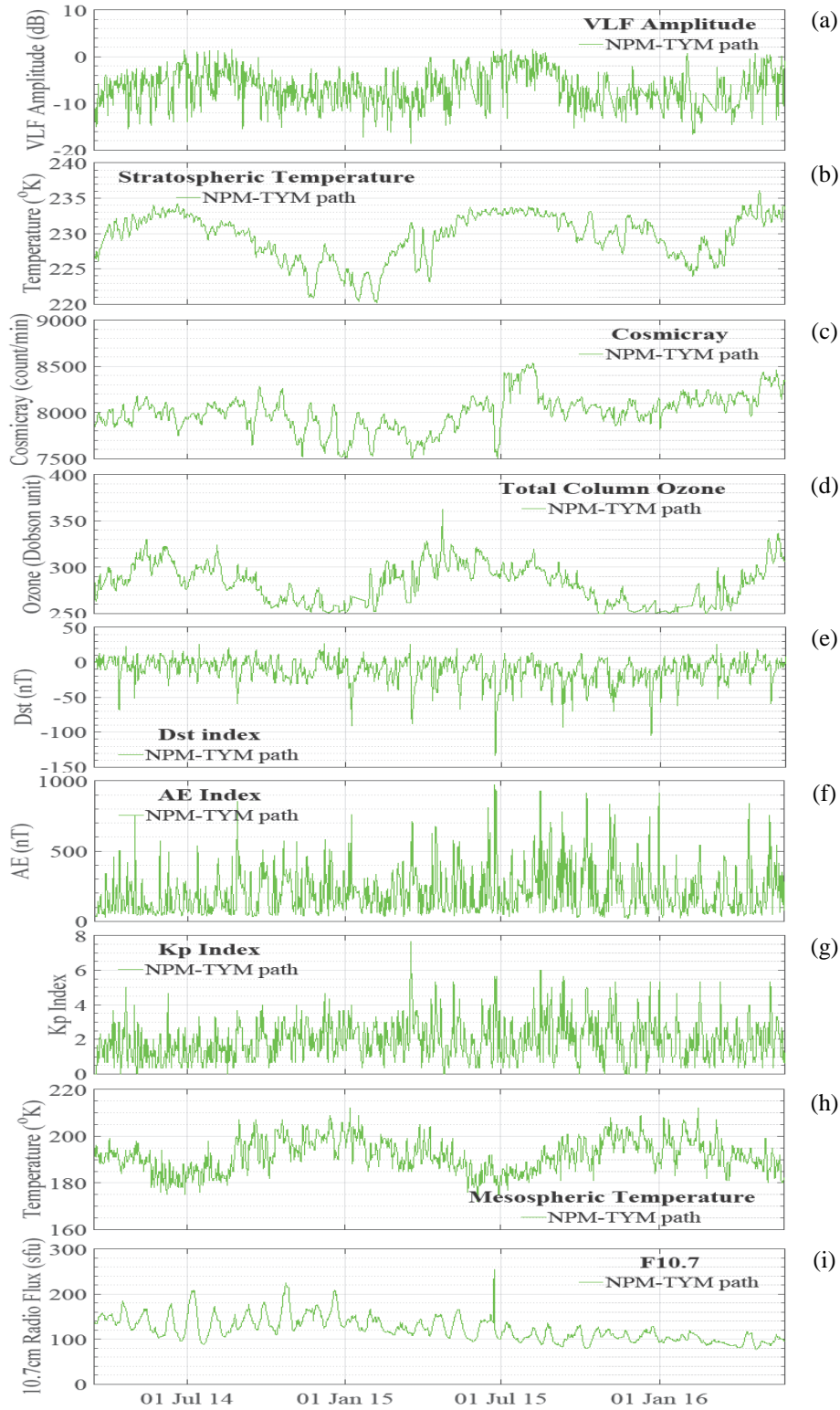


Figure 5.4: Daily nighttime mean values of the target and inputs of NARXNN for mid-latitude path from Hawaii, USA (NPM) to Tsuyama, Japan (TYM) station over the time interval from 15 March 2014 to 26 May 2016. (a) VLF daily nighttime mean amplitude for NPM-TYM path as an input and target of the system (dB), and seven parameters as inputs to the system, (b) stratospheric temperature ($^{\circ}$ K), (c) cosmic ray (in count/ min), (d) total column ozone (Dobson unit), (e) Dst index (nT), (f) AE index (nT), (g) Kp index, (h) mesospheric temperature ($^{\circ}$ K), and (i) F10.7 index (sfu).

Figure 5.4 shows daily nighttime mean values of the target and inputs of NARXNN for the mid-latitude (NPM-TYM) path over the time interval from 15 March 2014 to 26 May 2016. This figure is derived from the raw data with range time between 12 UT and 14 UT. Figure 5.4(a) illustrates the nighttime VLF amplitude received at Tsuyama VLF receiver station over three years from NPM VLF transmitter station at Pearl Harbour, Lualuahei, Hawaii, United States of America. Figure 5.4(b) depicts the daily mean of the nighttime stratospheric temperature time series obtained from the Atmospheric Infrared Sounder (AIRS) instrument on the NASA's Earth Observing System (EOS) Aqua satellite measuring during the orbits of descending node. Figure 5.4(c) shows the daily intensity of the cosmic ray time series data was calculated based on the time series of count/min from the Magadan cosmic ray station operated by the Institute of Cosmophysical Research and Radio Propagation Wave, part of the Russian Academy of Sciences. The daily mean of total column ozone (TCO) data in Dobson unit was obtained from the Ozone Monitoring Instrument (OMI) aboard on the NASA's AURA spacecraft is shown Figure 5.4(d). The Dst, AE and Kp indices are shown in Figure 5.4(e), Figure 5.4(f) and Figure 5.4(g) respectively obtained from the World Data Center (WDC) for geomagnetism, Kyoto, Japan. The daily mean mesospheric temperature data in Kelvin at height of 80-90 km was obtained from the SABER on the NASA satellite is shown Figure 5.4(h). Finally, Figure 5.4(i) shows the time series of F10.7 index obtained from the Interplanetary Magnetic Field (IMF) and plasma data.

5.1.3 High-Latitude VLF Propagation Path

Figure 5.5 illustrates daily nighttime mean values of the target and inputs of NARXNN for the high-latitude (NLK-CHF) path over the time interval from 1 January 2011 to 31 December 2013. This figure is derived from the raw data with range time between 10:30 UT and 11:30 UT. Figure 5.5(a) shows the nighttime VLF amplitude received at CHF over three years from NLK VLF transmitter station at Oso Wash, Jim Creek, Washington, United States of America. Figure 5.5(b) shows the daily mean of the nighttime stratospheric temperature time series obtained from the AIRS instrument measuring during the orbits of descending node. Figure 5.5(c) shows the daily intensity of the cosmic ray counts as time series. The daily mean of total column ozone (TCO) is shown Figure 5.5(d). Furthermore, Figure 5.5(e), Figure 5.5(f) and Figure 5.5(g) show the time series of Dst, AE and Kp indices respectively. The daily mean mesospheric temperature is shown Figure 5.5(h). Finally, Figure 5.5(i) shows the time series of F10.7 index.

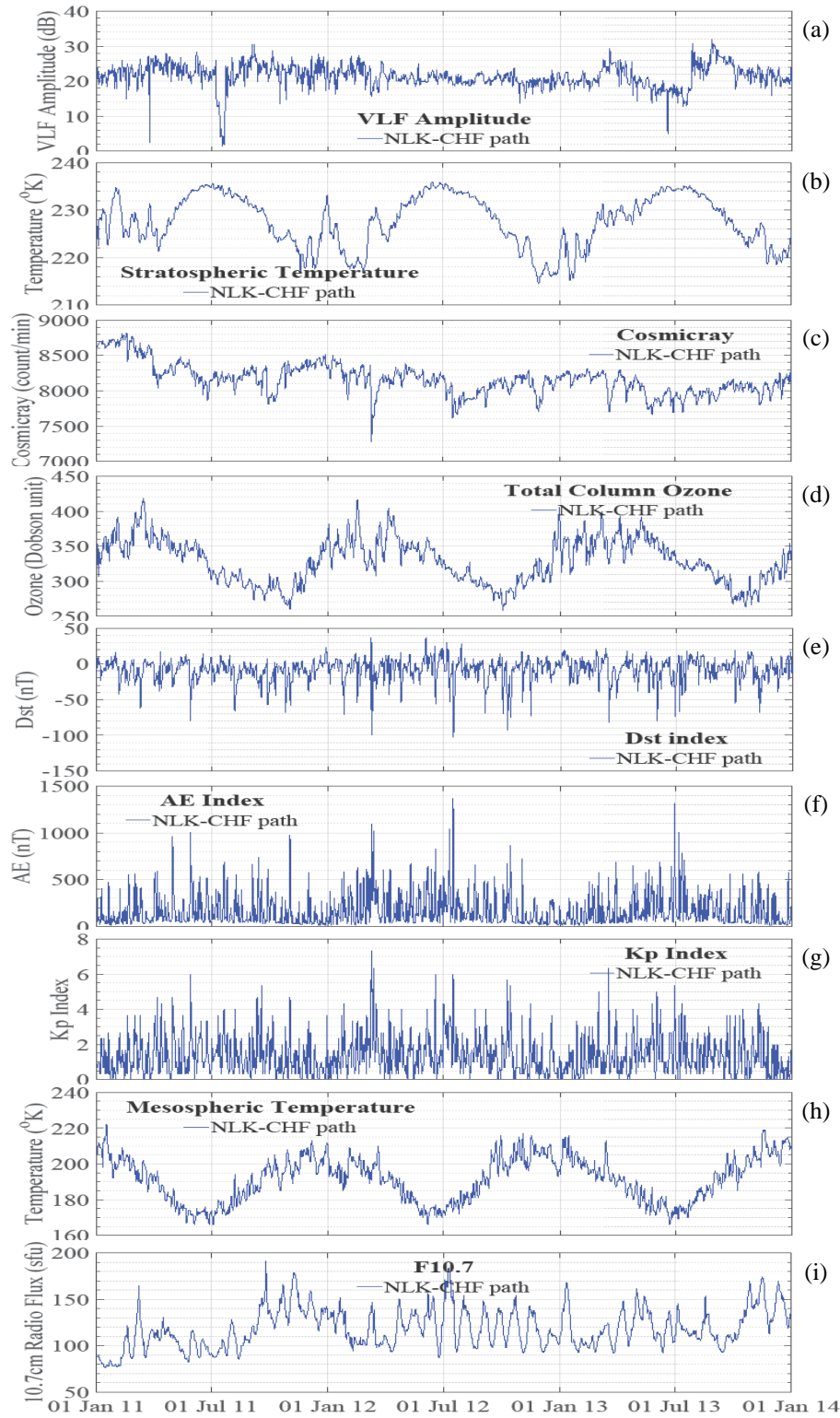


Figure 5.5: Daily nighttime mean values of the target and inputs of NARXNN for low-mid-latitude path from Washington, USA (NLK) to Chofu, Japan (CHF) station over the time interval from 1 January 2011 to 31 December 2013. (a) VLF daily nighttime mean amplitude for NLK-CHF path as an input and target of the system (dB), and seven parameters as inputs to the system, (b) stratospheric temperature ($^{\circ}\text{K}$), (c) cosmic ray (in count/ min), (d) total column ozone (Dobson unit), ((e) Dst index (nT), (f) AE index (nT), (g) Kp index, (h) mesospheric temperature ($^{\circ}\text{K}$), and (i) F10.7 index (sfu).

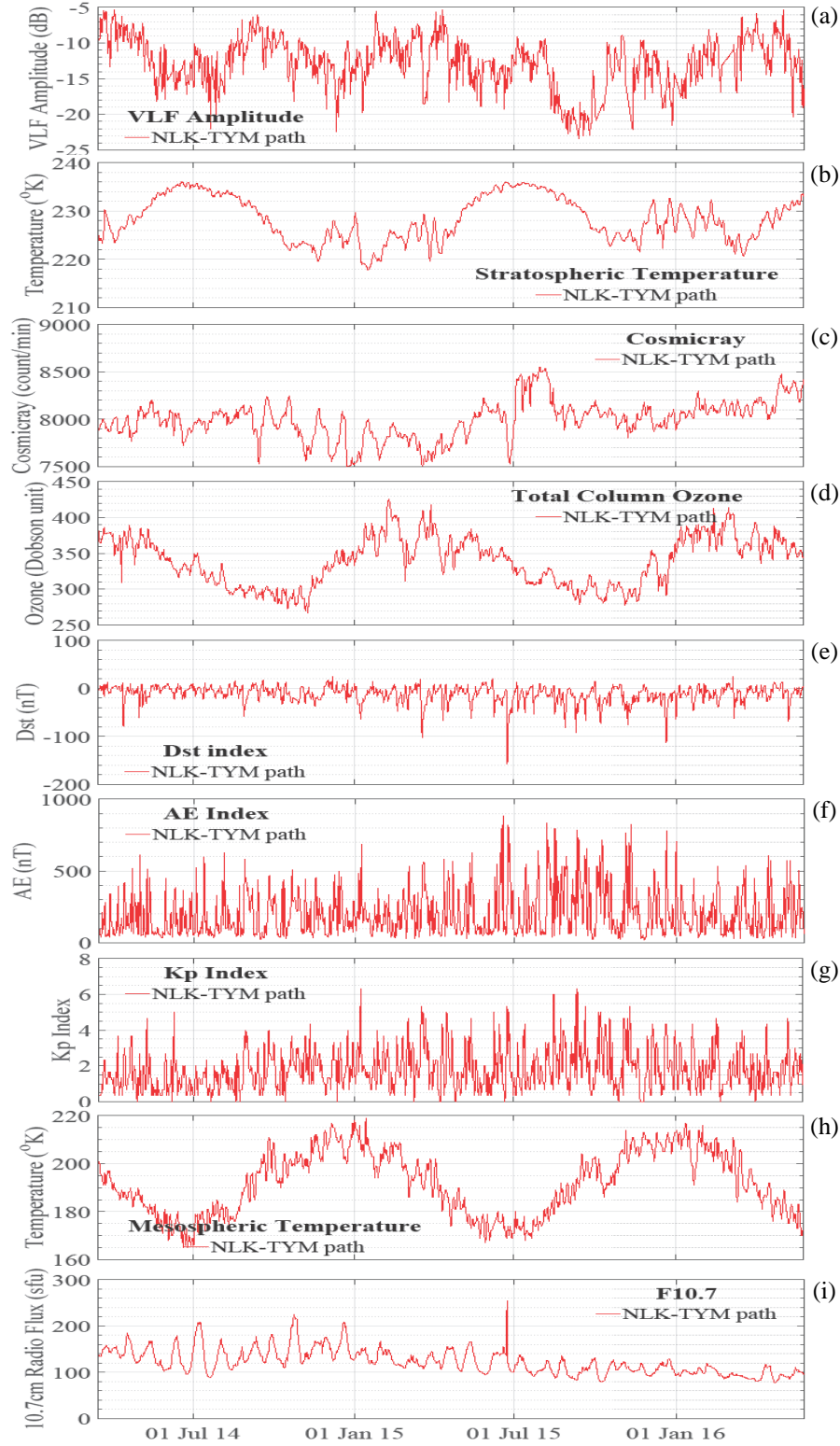


Figure 5.6: Daily nighttime mean values of the target and inputs of NARXNN for low-mid-latitude path from Washington, USA (NLK) to Tsuyama, Japan (TYM) station over the time interval from 15 March 2014 to 26 May 2016. (a) VLF daily nighttime mean amplitude for NLK-TYM path as an input and target of the system (dB), and seven parameters as inputs to the system, (b) stratospheric temperature ($^{\circ}\text{K}$), (c) cosmic ray (in count/ min), (d) total column ozone (Dobson unit), (e) Dst index (nT), (f) AE index (nT), (g) Kp index, (h) mesospheric temperature ($^{\circ}\text{K}$), and (i) F10.7 index (sfu).

Figure 5.6 shows daily nighttime mean values of the target and inputs of NARXNN for the high-latitude (NLK-TYM) path over the time interval from 15 March 2014 to 26 May 2016. This figure is derived from the raw data with range time between 10:30 UT and 11:30 UT. Figure 5.6(a) shows the nighttime VLF amplitude received at Tsuyama VLF receiver station over three years from NLK VLF transmitter station at Oso Wash, Jim Creek, Washington, United States of America. This variation appears to have some local maxima and minima for summer and winter times respectively in the northern hemisphere, however it indicates a complex response. Figure 5.6(b) shows the daily mean of the nighttime stratospheric temperature time series obtained from the AIRS instrument measuring during the orbits of descending node. Figure 5.6(c) shows the daily intensity of the cosmic ray counts as time series. The daily mean of total column ozone (TCO) is shown Figure 5.6(d). Furthermore, Figure 5.6(e), Figure 5.6(f) and Figure 5.6(g) show the time series of Dst, AE and Kp indices respectively. The daily mean mesospheric temperature is shown Figure 5.6(h). Finally, Figure 5.6(i) shows the time series of F10.7 index. As seen from Figure 5.6, the temporal dependences of the TCO, stratospheric and mesospheric temperatures appear to have similar seasonal dependences but with some timing offset, while the cosmic ray has a systematic descending trend over three years presumably due to modulation from changing solar activity.

5.2 One-step Ahead Prediction (OSA)

In this thesis, based on the results in Section 5.4 which the NARXNN with 3 delay time and 200 neurons using LMANN algorithm was the best nonlinear dynamic model to predict the daily nighttime mean value of VLF electric field amplitude. The first datasets for OSA training was chosen from 1 to 766 days of the total data (1096 days) and for validation was used the remaining dataset (767 – 1096 days) over the time interval from 1 January 2011 to 31 December 2013 and second datasets for OSA training was chosen from 1 to 563 days of the total data (804 days) and for validation was used the remaining dataset (564 – 804 days) over the time interval from 15 March 2014 to 26 May 2016. For the learning purpose, the network training function that updates the weight and bias values according to Levenberg-Marquardt optimization was modified to include the regularization technique. It minimizes a combination of squared errors and weights, then the correct combination to produce a network which generalizes well. The training was stopped when the lowest RMSE between each data set and its prediction were reached. The early stopping technique used In this thesis involves

simultaneous training and validation. Further, this technique to prevent overfitting during the network training.

The NARXNN OSA constructed model was used to predict three different latitude paths namely low-mid-latitude path (NWC-CHF and NWC-TYM), middle-latitude path (NPM-CHF and NPM-TYM), and high-latitude path (NLK-CHF and NLK-TYM).

5.2.1 Architecture of NARXNN OSA Modeling

The constructed network consists of three main layers namely input layer, hidden layer, and output layer. First, the input layer is representing of the current and previous stage of inputs and outputs. And then, the input layer value will be fed into the hidden layer. Second, the hidden layer is representing of one or few number of neurons resulting in a nonlinear mapping of the connected combination of the weights values from the input layer. Third, the output layer is representing of the connected combination of the weights values from the hidden layer. The dynamical order of inputs and outputs and number of neurons in each layer are pre-determined. *Haykin* [1998] is represent a few methods for determining of those values. An appropriate training algorithm and performance measure has to be selected. Finally, the category of the nonlinear map also requires to be determined.

The input and target values need for pre- and post-processing in order to obtain a valid training set. Further, those processes consist of normalization of the input and target data into values with the interval from -1 to 1 , normalization values for the inputs and targets to obtain zero mean and unity variance and removal of constant inputs and outputs. For evaluating the performance of a network, performance criteria are chosen to observe the error between desired responses (original data) and calculated outputs (prediction).

In this thesis, the NARXNN, where the exogenous inputs (u) are stratospheric temperature, cosmic rays, total column ozone, dst index, AE index, Kp index, mesospheric temperature, and F10.7 have been implemented using Levenberg Marquardt Neural Network (LMANN) Algorithm. The VLF electric amplitude field was used as the output (y). Both inputs and output have a memory of 3 days from the given day. The NARXNN consist 200 neurones in the hidden layer each utilizing tangent sigmoid activation function. The NARXNN model for one-step ahead prediction has been diagrammatically shown in Figure 5.7. Further, the NARXNN model for the temporal dependence of the VLF amplitude can be given by equation (5.1).

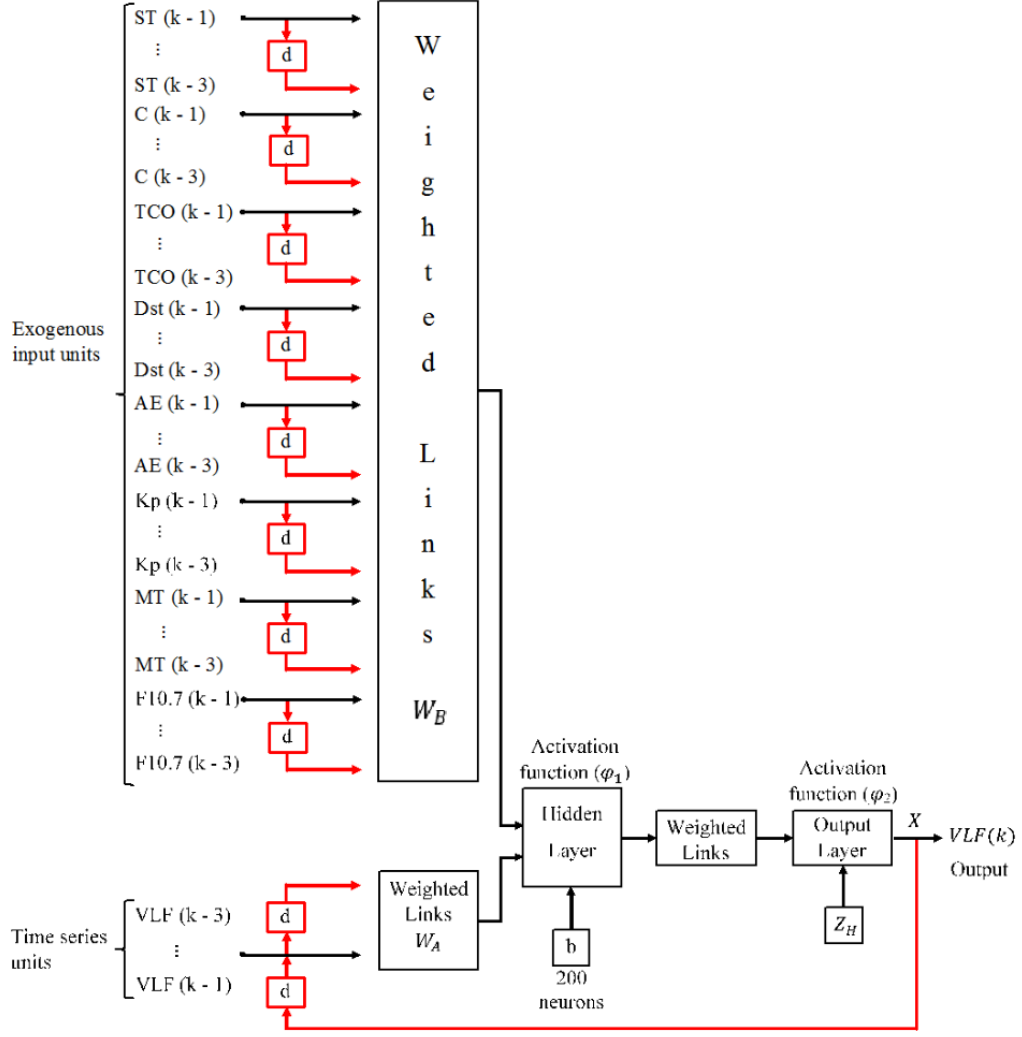


Figure 5.7: NARXNN architecture of the VLF amplitude prediction.

$$\begin{aligned}
 VLF(k+1) = F[& (VLF(k), \dots, VLF(k-d_{VLF} \\
 & +1); ST(k), \dots, ST(k-d_T+1); C(k), \dots, C(k \\
 & -d_C+1); TCO(k), \dots, TCO(k-d_{TCO} \\
 & +1); Dst(k), \dots, Dst(k-d_{Dst} \\
 & +1); AE(k), \dots, AE(k-d_{Kp} \\
 & +1); Kp(k), \dots, Kp(k-d_{Kp} \\
 & +1); MT(k), \dots, MT(k-d_{Kp} \\
 & +1); F10.7(k), \dots, F10.7(k-d_{Kp}+1)]
 \end{aligned} \tag{5.1}$$

where $F[\cdot]$ denotes a nonlinear tangent sigmoid function and d_{VLF} , d_{ST} , d_C , d_{TCO} , d_{Dst} , d_{AE} , d_{Kp} , d_{MT} and $d_{F10.7}$ are maximum lagged time in step VLF for the output, and stratospheric temperature, cosmic ray, total column ozone, Dst index, AE index, Kp index, mesospheric

temperature and F10.7 for inputs, respectively. Moreover, nonlinear tangent sigmoid function given by equation (5.2)

$$F(x) = \frac{1}{1 + \exp(-x)} - \frac{1}{2} \quad (5.2)$$

Furthermore, to calculate the coefficient of the NARXNN model in equation (5.1), we used Garson's algorithm [Garson, 1991] with modified by Goh [1995] for partitioning the neural network connection weights in order to determine the relative importance of each input variable in the network as represented in equation (5.3). Based on the training process, we able to the interpretation of connection weights and variable contributions in neural networks. Next, we illustrating how connection weights and the overall influence of the input variables in the network can be assessed statistically.

$$RI_x = \sum_{h=1}^n \frac{|w_{uh}w_{hy}|}{\sum_{x=1}^m |w_{uh}w_{hy}|} \quad (5.3)$$

where RI is relative importance of input variable in the network, x is input of the network, h is hidden layer, y is output, w_{uh} is weight between input and hidden layer neurons, and w_{hy} is weight between hidden layer and output neurons.

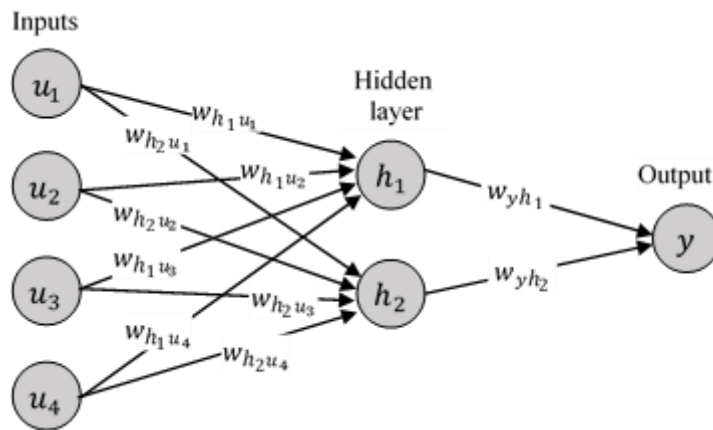


Figure 5.8: Calculating relative importance of each input variable using the Garson's algorithm.

Moreover, the example of relative importance based on Figure 5.8. First, the contribution of each input neuron to output via each hidden neuron calculated as the product of the input-hidden and hidden-output connection as described:

$$C_{h_1u_1} = w_{h_1u_1} \times w_{yh_1} \quad (5.4)$$

next, the relative contribution of each neuron to the outgoing signal of each hidden neuron.

$$r_{h_1u_1} = \frac{|C_{h_1u_1}|}{(|C_{h_1u_1} + C_{h_1u_1} + C_{h_1u_1} + C_{h_1u_1}|)} \quad (5.5)$$

then, sum of the input neuron contribution

$$S_1 = r_{h_1u_1} + r_{h_2u_1} \quad (5.6)$$

finally, relative importance (RI) of each input variable

$$RI_1 = \frac{S_1}{(S_1 + S_2 + S_3 + S_4)} \quad (5.7)$$

The NARXNN performances can be evaluated using root mean square error (RMSE). Further, the RMSE can range between 0 and ∞ and it should be noted that the minimum RMSE value is the best network performance in prediction. In order to calculate the RSME, the difference between estimated and actual values are each squared and then averaged over the sample period and then rooted. RMSE can be given by equation (5.8)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (error_{sim}(i))^2} \quad (5.8)$$

where n is the number of data points (one point per day), and $error_{sim}(i)$ is the error difference (in dB) between the VLF amplitude from observation and the model prediction at the certain data point (day i).

The Pearson's correlation coefficient (CC) is also used to provide information about the NARXNN performances. Further, the CC can take a range of values from +1 to -1. A value of 0 indicates that there is no association between the two variables. A value greater than 0 indicates a positive association; that is, as the value of one variable increases, so does the value of the other variable. A value less than 0 indicates a negative association; that is, as the value of one variable increases, the value of the other variable decreases. The CC of prediction data over the observed values is calculated using (5.9)

$$CC = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}} \quad (5.9)$$

where

- N the total number of predicted value
- x_i the observed data
- y_i the predicted value
- \bar{x} the mean value of x_i
- \bar{y} the mean value of y_i

5.2.2 Prediction Results

Once the data sets (ordered points) are selected, they are used toward training a suitable NARXNN nonlinear dynamic model and then validating it. To achieve prediction model and to be able to identify the dynamic changing of the daily nighttime value of VLF electric field amplitude at a certain point in time in future, the time evolution of the input datasets has to be learned. The task in the training phase is to obtain optimal values for the adjustable parameters, weights, and biases, necessary to achieve the best fit between input and output data. This shows the ability of the trained NARXNN constructed model to dynamically map the historical and current data into the future. When the network is trained properly, we use the NARXNN model equation to build OSA model predictor.

In the following sections, we first demonstrate different latitude of VLF wave propagation path and also two different receiving station for a comparative study. By following

these cases the selection of the data and the network parameters become clearer and how the results vary by changing the latitude path and receiving station. We will compare the actual and the predicted values which are the NARXNN outputs. We also study them statistically and depict the errors.

5.2.2.1 Low-mid-latitude VLF propagation path

The NARXNN equation shows the significant parameters which are influenced to the constructed model as represented with the coefficients. The small coefficient means the contributions to the prediction of VLF amplitude also small and vice versa [Goh, 1995]. Equation (5.10) shows the complete NARXNN model equation for predicting VLF amplitude. Further, the first four significant terms for NWC-CHF path are described here. The first influential factor is the VLF electric field amplitude with 1 day before the given day $\{VLF(k-1)\}$ as represented by the coefficient of 1.431. The solar radio flux F10.7 with 2 days before the given day $\{F10.7(k-2)\}$ becomes the second significant factor with the coefficient of 1.330. The third influential factor is the Dst index with 2 days $\{Dst(k-2)\}$ before the given days with a weighting coefficient of 1.125. The last one is Kp index with 1 day before the given day $\{Kp(k-1)\}$ as indicated by the coefficient of 0.977.

$$\begin{aligned}
 VLF(k) = F [& -1.431VLF(k-1) + 0.230VLF(k-2) \\
 & + 0.378VLF(k-3) + 0.098ST(k-1) \\
 & - 0.643ST(k-2) - 0.577ST(k-3) \\
 & - 0.100C(k-1) - 0.818C(k-2) \\
 & + 0.899C(k-3) + 0.277TCO(k-1) \\
 & + 0.455TCO(k-2) - 0.350TCO(k-3) \\
 & + 0.430Dst(k-1) + 1.125Dst(k-2) \\
 & - 0.644Dst(k-3) - 0.489AE(k-1) \\
 & - 0.146AE(k-2) + 0.376AE(k-3) \\
 & + 0.977Kp(k-1) + 0.969Kp(k-2) \\
 & - 0.930Kp(k-3) + 0.418MT(k-1) \\
 & - 0.160MT(k-2) - 0.755MT(k-3) \\
 & + 0.905F10.7(k-1) - 1.330F10.7(k-2) \\
 & + 0.335F10.7(k-3) + 1.993]
 \end{aligned} \tag{5.10}$$

Moreover, the TYM station NARXNN model equation can be represented in equation (5.11) and for the first four significant terms for NWC-TYM path are described here. The first influential factor is the VLF electric field amplitude with 1 day before the given day $\{VLF(k-1)\}$ as represented by the coefficient of 1.6817. The solar radio flux F10.7 with 1 day before the given day $\{F10.7(k-1)\}$ becomes the second significant factor with the coefficient of 1.1671. The third influential factor is the Dst index with 1 day before the given days $\{Dst(k-1)\}$ with a weighting coefficient of 0.9754. The last one is Kp index with 1 day before the given day $\{Kp(k-1)\}$ as indicated by the coefficient of 0.977.

$$\begin{aligned}
 VLF(k) = F [& -1.682VLF(k-1) + 0.688VLF(k-2) \\
 & - 0.368VLF(k-3) + 0.248ST(k-1) \\
 & - 0.871ST(k-2) - 0.722ST(k-3) \\
 & - 0.042C(k-1) - 0.112C(k-2) \\
 & + 0.133C(k-3) + 0.081TCO(k-1) \\
 & - 0.395TCO(k-2) + 0.200TCO(k-3) \\
 & + 0.975Dst(k-1) + 0.288Dst(k-2) \\
 & - 0.262Dst(k-3) - 0.090AE(k-1) \\
 & + 0.039AE(k-2) - 0.002AE(k-3) \\
 & + 0.950Kp(k-1) + 0.414Kp(k-2) \\
 & - 0.401Kp(k-3) + 0.455MT(k-1) \\
 & - 0.155MT(k-2) + 0.054MT(k-3) \\
 & - 1.167F10.7(k-1) + 0.297F10.7(k-2) \\
 & - 0.170F10.7(k-3) + 0.261]
 \end{aligned} \tag{5.11}$$

Figure 5.9 and Figure 5.10 show the most relative significant factor which has influences in the model predictor. Focusing on the most four relative significant parameter, both the receiving station between Chofu and Tsuyama have the same variable of VLF electric field amplitude one day before the given day as the first contributed parameters. Tsuyama station has a bigger percentage compare with Chofu station and it shows with the value of 8.81% and 14.49% for NWC-CHF and NWC-TYM respectively. Further, the second relative significant factor, for Chofu station is solar radio flux at 10.7 cm (F10.7) index with two days before the given day as represented in Figure 5.9 with the value of 8.19% and Tsuyama station is F10.7 index with one day before the given day as illustrated in Figure 5.10 with the value of

10.37%. This difference in the delay time in the low-mid-latitude path may be due to unknown parameter outside the considered inputs. Moreover, the third relative significant parameter also different between NWC-CHF and NWC-TYM paths as shown Dst index two days before the given day with value of 6.93% and one day before the given day with value of 8.66% respectively. Finally, the fourth relative significant factor has the same parameter in both receiving stations with Kp index one day before the given day and the value of 6.01% and 8.44% as depicted in Figure 5.9 and Figure 5.10 respectively. NWC-TYM has a similar percentage as NWC-CHF with the various value of 0.08%.

Further, VLF amplitude has a higher contribution in the prediction model, it indicates that variability of VLF amplitude influence the model prediction in the same day show there is no strong external forcing exist or too complex disturbances between low to mid-latitudes path. Solar activity index (F10.7) is recognized as the most significant external forcing for VLF prediction model. The increasing long-term trend of the atomic concentration in the low region around 87-95 km at low-mid-latitude with the airglow measurement shows good agreement with the variation of F10.7 flux [*Clemesha et al.*, 2005]. The reason is the rate limiting reaction in the production of excited hydroxyl through the hydrogen-ozone mechanism and produce ozone from $O + O_2 + M$ to $O_3 + M$ and one of possible transport agent is eddy diffusion. Moreover, *Pakhomov and Gorbunov* [1983] was shown a positive correlation coefficient of 0.77 between F10.7 index and electron density at altitude 75 km based on 14 rockets probe measurement of electron density at Thumba with latitude of 8° N. Further, the variation of ionospheric dynamo electric field is induced by the global scale wind or the conductivity distribution. The effects of the variation of Cowling conductivity to the eastward electric field based on the observation data and the numerical model respectively. The increase of Cowling conductivity due to increase of solar activity changes the ionospheric dynamo electric field and further results in the weakening of eastward electric field and the decrease of the upward vertical $E \times B$ drift velocity.

The geomagnetic activity indices are the third and fourth significant parameters which are contributed to the model. *Kumar et al.* [2015] was mentioned the geomagnetic storm at low-latitude produced a significant reduction in the VLF amplitude around 3.2 dB to the average value of five quiet days. Moreover, storm effects on the D-region ionosphere was indicated that VLF low-latitude path (NWC to Suva, Fiji island) affected significantly up to 46 hours or near two-day. The duration is long but this shorter than storm effects a high-latitude which occurred for several days [*Kleimenova et al.*, 2004]. The electron density changing would have caused additional attenuation resulting in a substantial decrease in the signal

strength due to storm-induced D-region compositional changes. One of the possible mechanisms could be prompt penetration of the high-latitude electric fields to low-latitudes during the main phase and the electric fields generated by the disturbance dynamo. In the VLF low-mid-latitude path has not a high effect from the geomagnetic activity compare to high-latitude path because of the occurrence rate of anomalies around 31% in 2012 [Tatsuta *et al.*, 2015]. Furthermore, the relationship of the low-mid-latitude VLF radio emission and Kp index indicate a good correlation, such that the VLF emission exhibit a slight increase with Kp index [Hayakawa, 1989].

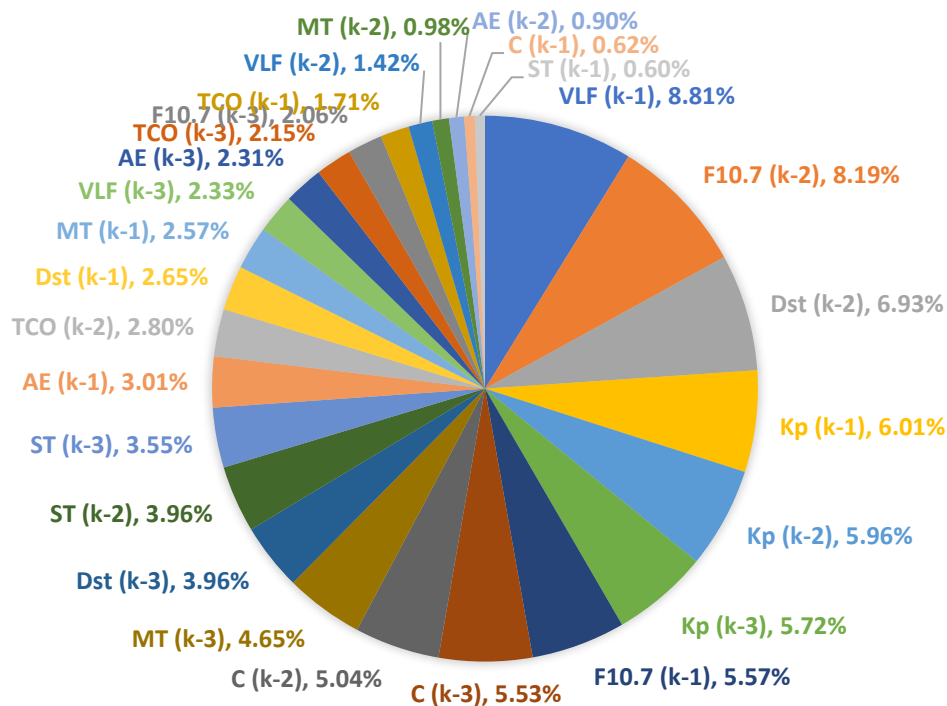


Figure 5.9: The relative significant parameter in low-mid-latitude NARXNN model of the daily nighttime mean values of VLF electric field amplitude Chofu station.

The NARXNN model with 3 days of input-memory and 200 neurons in the hidden layer using LMANN algorithm is used to predict the daily nighttime of VLF electric field amplitude. The OSA prediction results for low-mid-latitude path represented by Figure 5.11. The fitted model (inside the training period) with the time interval from 1 January 2011 to 4 February 2013 show in red curve and observation in blue curve. As a result, the fitted model has a good agreement with the original data for the low-mid-latitude path. The constructed model worked well for prediction as represented by Pearson correlation coefficient (r) is high and RMSE is small. The correlation coefficient for NWC-CHF path 0.913 and RMSE 1.50 dB.

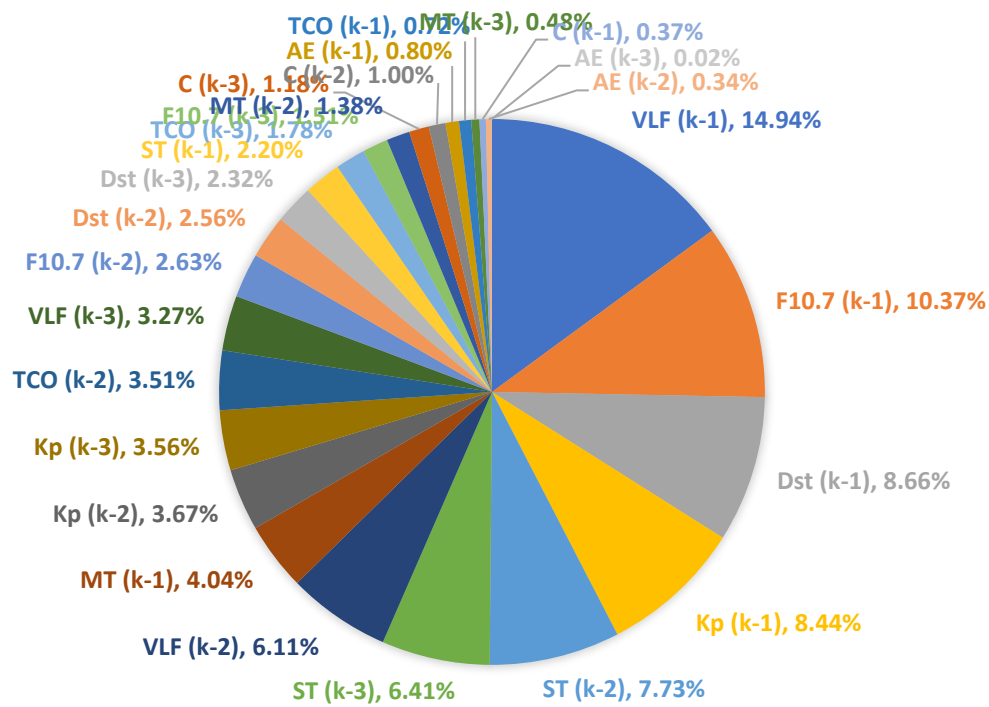


Figure 5.10: The relative significant parameter in low-mid-latitude NARXNN model of the daily nighttime mean values of VLF electric field amplitude Tsuyama station.

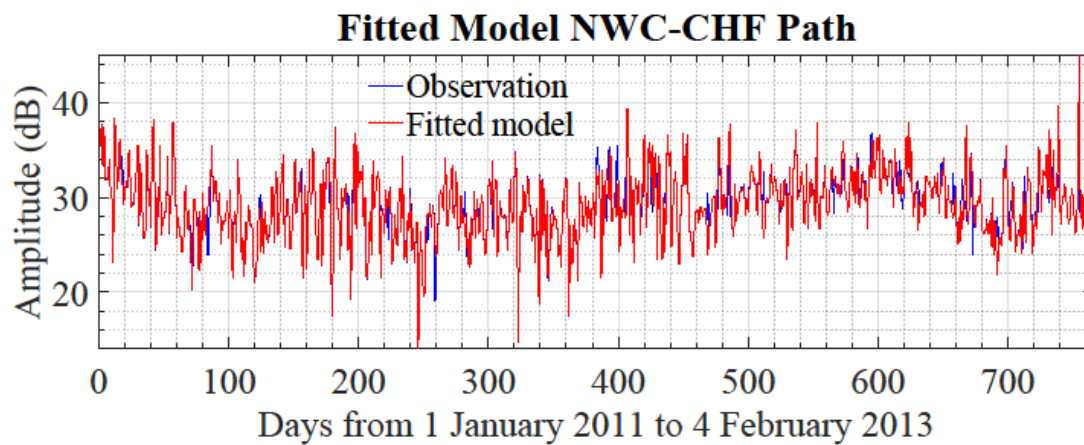


Figure 5.11: The fitted model predictions of NARX NN model of daily nighttime mean amplitude of VLF waves low-mid-latitude path with three-day of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 1 January 2011 to 4 February 2013 (VLF observation-blue solid; The fitted model-red dotted).

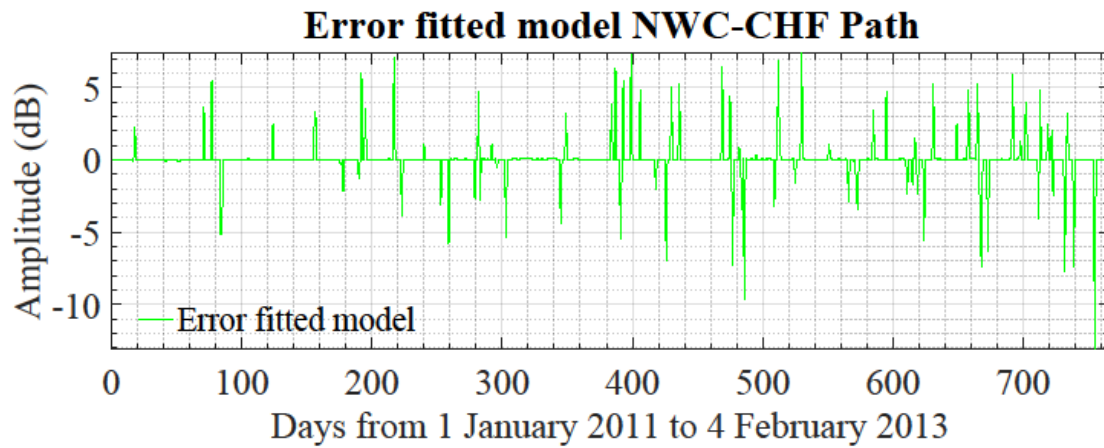


Figure 5.12: Error fitted model predictions of NARX NN model of daily nighttime mean amplitude of VLF waves low-mid-latitude path over the time interval from 1 January 2011 to 4 February 2013.

As seen from Figure 5.12, there are few dates with a relatively large error which mean big discrepancy between observed and fitted model values. This may be due to the physical factors other than we consider in the model inputs. Therefore, these observed values are considered to be anomalies due to unknown physical reason in the proposed model.

For comparative study and to examine the capability of our NARXNN model, we feed the constructed model with different datasets from different receiving station over the time interval between 15 March 2014 and 28 September 2015. The daily nighttime mean value of VLF electric field amplitude from Tsuyama station is predicted as the output and daily nighttime mean values of stratospheric and mesospheric temperatures, Dst, AE, Kp, and F10.7 indices, cosmic ray and total column ozone as the input. The results are depicted in Figure 5.13. The blue curve denotes the actual VLF amplitude values and the red curve shows the predicted ones. The fitted model also has a good agreement with Chofu station with the original data for the low-mid-latitude path. The NARXNN predictor model successful for prediction as represented by Pearson correlation coefficient (r) of 0.915.

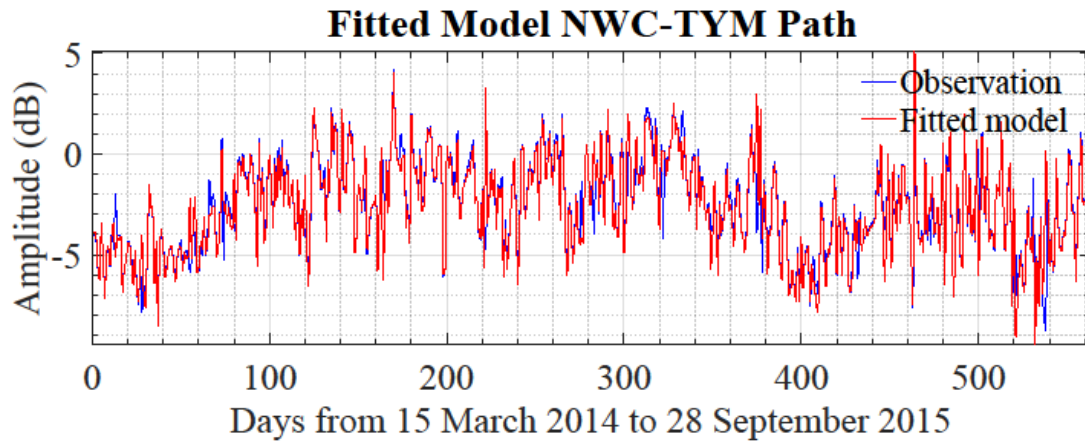


Figure 5.13: The fitted model predictions of NARX NN model of daily nighttime mean amplitude of VLF waves low-mid-latitude path with three-day of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 15 March 2014 to 28 September 2015 (VLF observation-blue solid; The fitted model-red dotted).

Furthermore, the prediction error for 563 days data point is shown graphically in Figure 5.14. The error varies from -8.9 dB to 5.6 dB and the RMSE is 0.98 dB. We have to consider a few dates with a relatively large error which mean big discrepancy between observed and fitted model values in low-mid-latitude Tsuyama station. This may be due to the physical factors other than we consider in the model inputs. Therefore, these observed values are considered to be anomalies due to unknown physical reason in the proposed model.

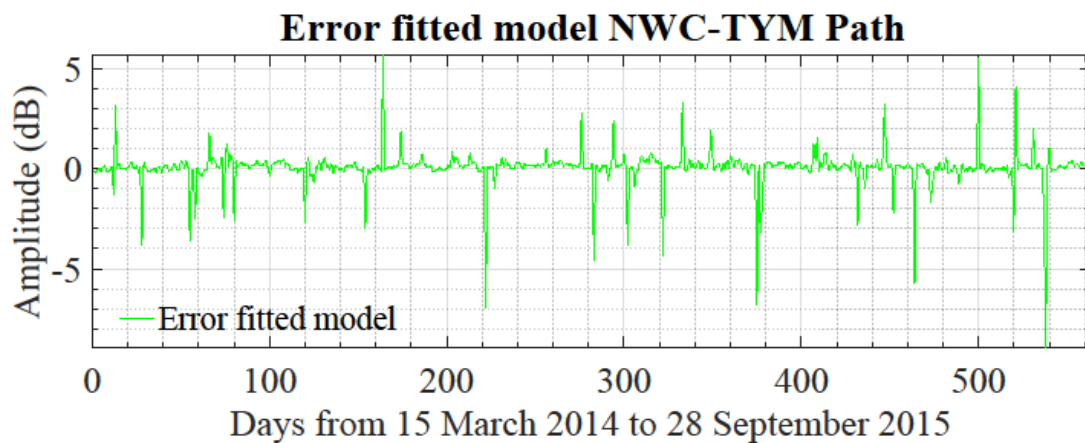


Figure 5.14: Error fitted model predictions of NARX NN model of daily nighttime mean amplitude of VLF waves low-mid-latitude path over the time interval from 15 March 2014 to 28 September 2015.

Figure 5.15 shows the OSA prediction with the data set outside the training period. The correlation coefficient for the prediction remains high value as represented by r of 0.910.

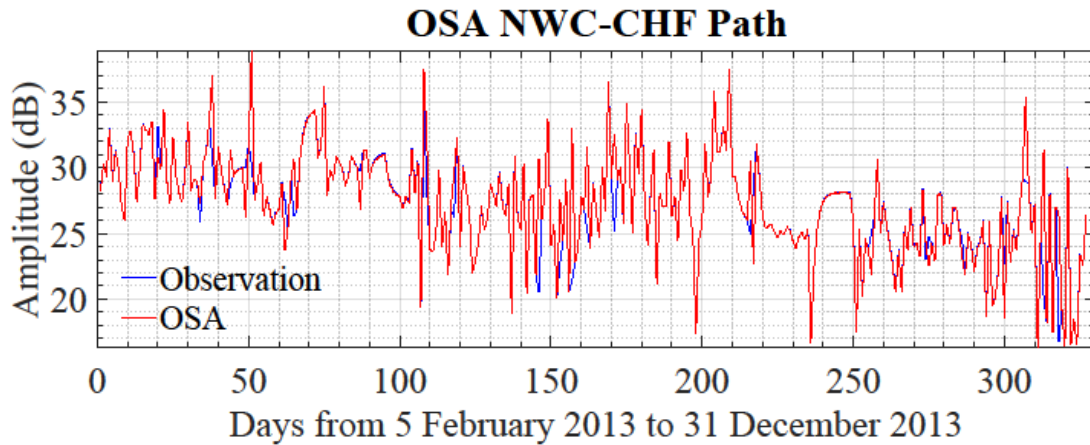


Figure 5.15: One Step (1 day) Ahead (OSA) predictions of NARX NN model of daily nighttime mean amplitude of VLF waves low-mid-latitude path with three-day of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 5 February 2013 to 31 December 2013 (VLF observation-blue solid; Prediction-red dotted).

Moreover, the prediction error for 330 days data point outside the training period is shown graphically in Figure 5.16. The error varies from -11.5 dB to 5.7 dB and the RMSE is 1.68 dB.

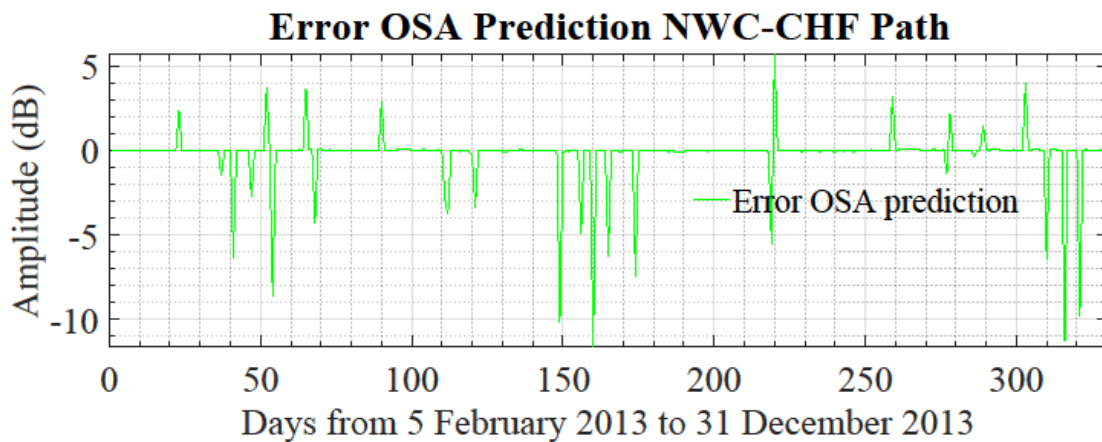


Figure 5.16: Error One Step (1 day) Ahead (OSA) predictions of NARX NN model of daily nighttime mean amplitude of VLF waves low-mid-latitude path over the time interval from 1 January 2011 to 4 February 2013.

Further, to examine the capability of our NARXNN model, we feed the built model with different datasets from different receiving station over the time interval between 29 September 2015 to 26 May 2016. We still use NARXNN model equation for low-mid-latitude path Tsuyama station with three days delay time $\{d_u = 3, d_y = 3\}$. The results are illustrated in Figure 5.17. The blue curve denotes the actual VLF amplitude values and the red curve

shows the predicted ones. The OSA still has a good agreement as a fitted model with the original data for low-mid-latitude path. The NARXNN predictor model successful for prediction outside the training period as shown by Pearson correlation coefficient (r) of 0.909.

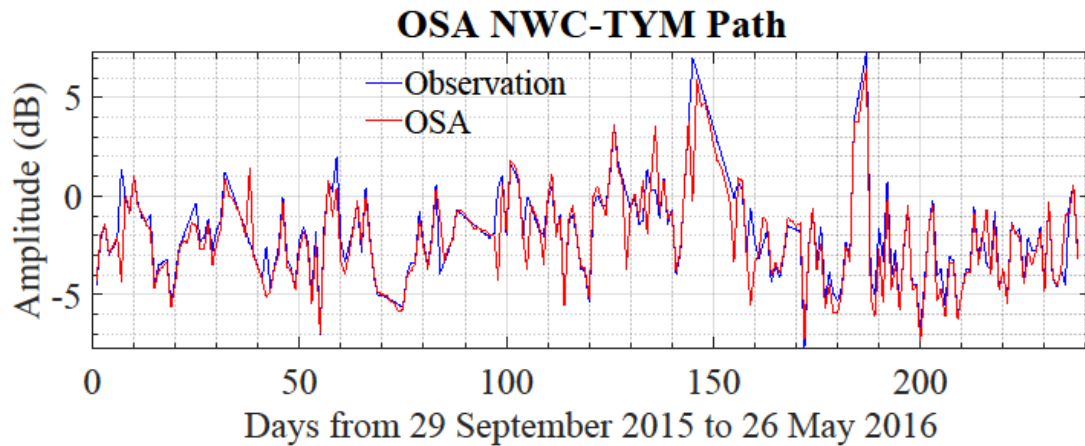


Figure 5.17: One Step (1 day) Ahead (OSA) predictions of NARX NN model of daily nighttime mean amplitude of VLF waves low-mid-latitude path with three-day of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 29 September 2015 to 26 May 2016 (VLF observation-blue solid; Prediction-red dotted).

Furthermore, the prediction error of OSA outside the training period for 241 days data point is shown graphically in Figure 5.18. The error varies from -3.7 dB to 7.2 dB and the RMSE is 1.1 dB. This result has a relatively good in average compared with Chofu station. This condition may be due to the physical factors other than we consider in the model inputs and also depend on the signals interferences on receiving site

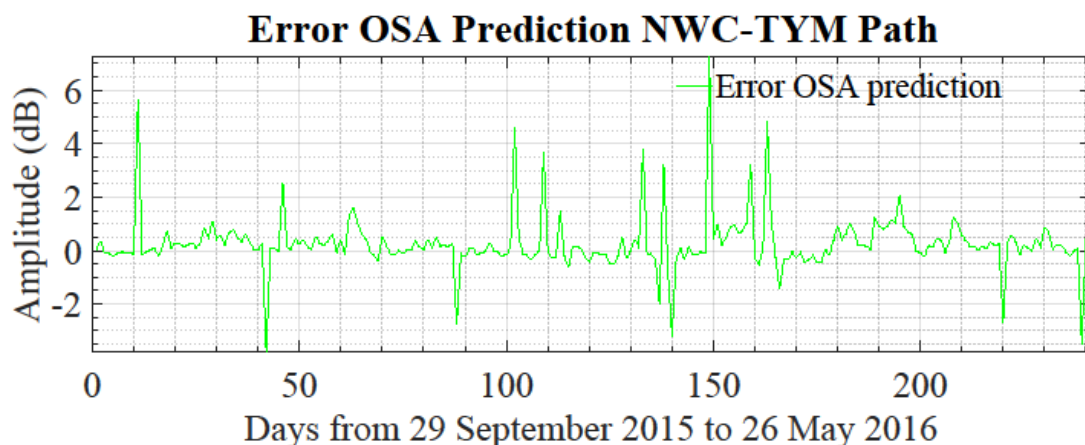


Figure 5.18: Error One Step (1 day) Ahead (OSA) predictions of NARX NN model of daily nighttime mean amplitude of VLF waves low-mid-latitude path over the time interval from 29 September 2015 to 26 May 2016.

5.2.2.2 Mid-latitude VLF propagation path

As seen in equation (5.12) that the stratospheric temperature with 2 days before the given day $\{ST(k - 2)\}$ becomes the first significant variable in NPM-CHF path as indicated by the coefficient of 2.272. The solar radio flux F10.7 with 1 day before the given day $\{F10.7(k - 1)\}$ is the second influential parameter with the coefficient of 2.190. The third influential factor is the VLF electric field amplitude with 1 day $\{VLF(k - 1)\}$ before the given day as indicated by the coefficient of 1.440. The fourth significant factor for NPM-CHF path is mesospheric temperature with 1 day before the given day $\{MT(k - 1)\}$ with the coefficient of 1.275.

$$\begin{aligned}
 VLF(k) = F [& -1.440VLF(k - 1) + 0.214VLF(k - 2) \\
 & - 1.087VLF(k - 3) - 0.591ST(k - 1) \\
 & - 2.272ST(k - 2) - 0.517ST(k - 3) \\
 & - 0.913C(k - 1) - 0.862C(k - 2) \\
 & - 0.502C(k - 3) - 0.688TCO(k - 1) \\
 & - 1.188TCO(k - 2) - 0.740TCO(k - 3) \\
 & + 1.153Dst(k - 1) - 0.470Dst(k - 2) \\
 & - 0.944Dst(k - 3) + 0.024AE(k - 1) \\
 & + 0.822AE(k - 2) + 0.915AE(k - 3) \\
 & - 0.018Kp(k - 1) - 0.018Kp(k - 2) \\
 & + 0.064Kp(k - 3) - 1.275MT(k - 1) \\
 & - 0.180MT(k - 2) + 0.512MT(k - 3) \\
 & - 2.190F10.7(k - 1) + 1.054F10.7(k - 2) \\
 & - 0.413F10.7(k - 3) - 2.183]
 \end{aligned} \tag{5.12}$$

Furthermore, the TYM station NARXNN model equation can be represented in equation (5.13) and for the first four significant terms for NPM-TYM path are described here. The first influential factor is the stratospheric temperature with two days before the given day $\{ST(k - 2)\}$ as represented by the coefficient of 2.371. The solar radio flux F10.7 with one day before the given day $\{F10.7(k - 1)\}$ becomes the second significant factor with the coefficient of 1.982. The third influential factor is the VLF electric field amplitude with one day before the given days $\{VLF(k - 1)\}$ with a weighting coefficient of 1.3061. Finally, the last parameter is mesospheric temperature with one day before the given day $\{MT(k - 1)\}$ as indicated by the coefficient of 1.201.

$$\begin{aligned}
VLF(k) = F [& 1.306VLF(k-1) - 0.418VLF(k-2) \\
& + 1.002VLF(k-3) + 0.896ST(k-1) \\
& + 2.371ST(k-2) + 0.904ST(k-3) \\
& - 0.937C(k-1) - 0.333C(k-2) \\
& - 0.397C(k-3) + 0.825TCO(k-1) \\
& - 1.067TCO(k-2) - 0.850TCO(k-3) \\
& + 1.060Dst(k-1) + 0.381Dst(k-2) \\
& + 0.054Dst(k-3) + 0.663AE(k-1) \\
& - 0.266AE(k-2) - 0.023AE(k-3) \\
& + 0.435Kp(k-1) + 0.423Kp(k-2) \\
& - 0.277Kp(k-3) - 1.201MT(k-1) \\
& + 0.280MT(k-2) - 0.121MT(k-3) \\
& + 1.982F10.7(k-1) - 0.971F10.7(k-2) \\
& + 0.370F10.7(k-3) + 0.507]
\end{aligned} \tag{5.13}$$

As represented in Figure 5.19 and Figure 5.20, the four of most relative significant factor which has influences in the model predictor on mid-latitude VLF propagation path. For the most relative significant parameter, both the receiving station between Chofu and Tsuyama have the same variable of stratospheric temperature two days before the given day. Tsuyama station has a bigger percentage compare with Chofu station and it shows with the value of 10.79% and 11.07% for NPM-CHF and NPM-TYM respectively. Further, the second relative significant factor, for both of the receiving stations are solar radio flux at 10.7 cm (F10.7) index with one day before the given day as represented in Figure 5.19 with the value of 10.40% and Tsuyama station is 10.00% as illustrated in Figure 5.20. Comparison of the second significant parameter Chofu station has bigger percentage than Tsuyama station. Moreover, the third relative significant parameter also same between NPM-CHF and NPM-TYM paths as shown VLF electric field amplitude one day before the given day with a value of 6.84% and 6.59% respectively. Finally, the fourth relative significant factor also has the same parameter in both receiving stations with mesospheric temperature one day before the given day and the value of 6.05% and 6.06% as depicted in Figure 5.19 and Figure 5.20 respectively. NPM-CHF has a similar percentage as NPM-TYM with the various value of 0.01%.

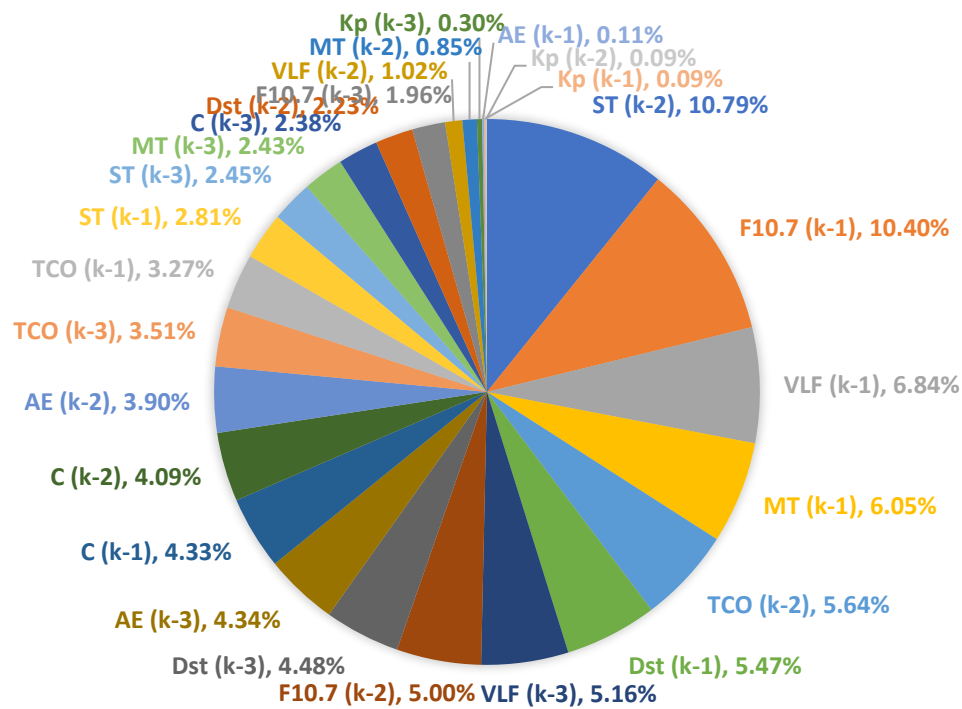


Figure 5.19: The relative significant parameter in mid-latitude NARXNN model of the daily nighttime mean values of VLF electric field amplitude Chofu station.

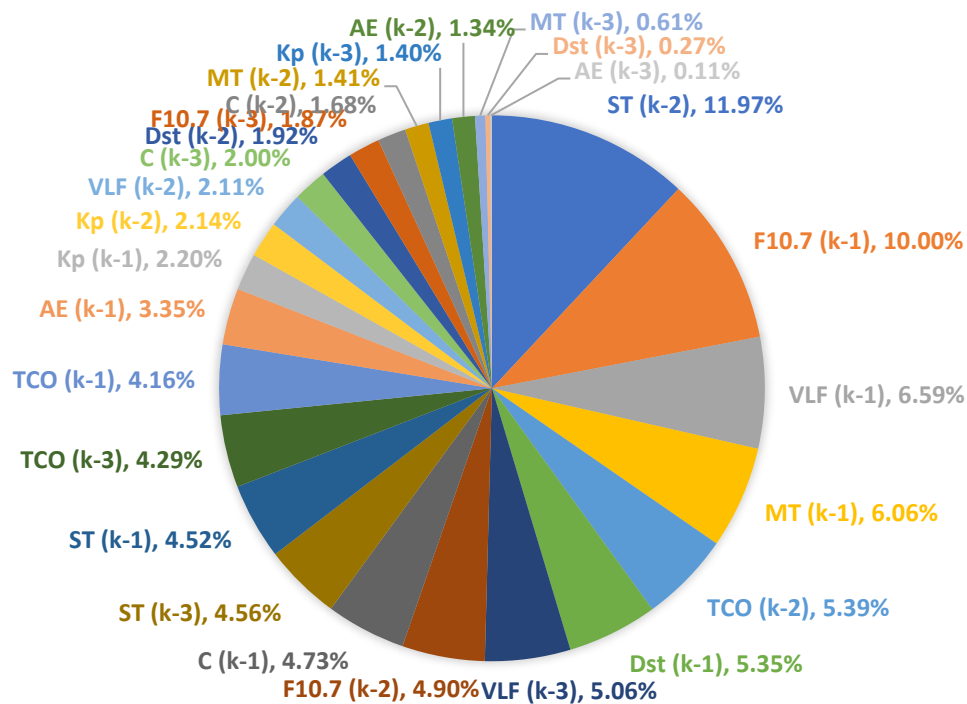


Figure 5.20: The relative significant parameter in mid-latitude NARXNN model of the daily nighttime mean values of VLF electric field amplitude Tsuyama station.

Stratospheric temperature is the most significant variable with the large coefficients that contributed in the prediction model. It indicates a good correlation between stratospheric temperature and the VLF amplitude in the mid-latitude path, as has previously been shown by *Pal and Hobara* [2016]. The vertical coupling is expected with the zonal direction because the observed stratospheric temperature is around the east-west VLF propagation path. Upward propagating gravity wave and/or a heat conduction wave may be possible agents transporting energy from stratospheric temperature disturbances to the lower ionosphere [*Volland and Warnecke*, 1968]. These waves have a propagation time for few days from the stratosphere to reach the lower ionosphere for nearly vertical incidence, which agrees with the observed time lag of 2-3 days.

The D-region electron concentration has dependence with solar activity which affected in the reflection of radio wave [*Danilov*, 1998]. The most obvious source of the preceding relation between D-region electron density variation with solar activity of the intensity of radiation which produces ionization or ionization rate. It is known that several sources are possible for D-region ionization such as the solar Lyman-alpha, which ionizes NO molecules, solar radiation in the interval of $\lambda = 111.8 - 102.7$ nm, which ionizes the excited $O_2(^1\Delta_g)$ molecules, solar X-rays below 3 nm, and galactic cosmic rays both of the last sources ionize all neutral molecules. Major source of the ionization in the nighttime D region is still recognized as being solar Lyman-alpha (121.6nm) scattered by the neutral hydrogen in the geocorona [*Banks and Kockarts*, 1973]. Further, variations of electron concentration in the D-region at an altitude of 90 km with F10.7 has a positive correlation with the correlation coefficient of 0.77 [*Pakhomov and Gorbunov*, 1983]. The variability of the VLF amplitude influence VLF amplitude prediction, which is highly correlated with the consecutive day. This high correlation indicates that the VLF amplitude does not vary a great deal from the value of the day before if no strong external forcing exists.

The VLF amplitude signal in the middle-latitude path and mesospheric temperature data from SABER instrument around 80-90 km altitude was found to has a very high negative correlation. Moreover, VLF amplitude decrease usually comes hand-in-hand with temperature increase in the mesosphere region [*Silber et al.*, 2013]. Modal interference has an important effect on the received amplitude of VLF signal, as the received amplitude is the sum of several waveguide modes [*Wait*, 1957]. The signal might be strongly attenuated if destructive modal interference occurs, which depends on the waveguide's characteristics and the distance between the transmitter and receiver. On the other hand, VLF absorption might also play a role in the observed day to day correlation between mesospheric temperature and VLF amplitudes.

The absorption factor is highly dependence with electron temperature in upper mesosphere. The modal interference and signal absorption are potentially capable to explain at least quantitatively the connection between VLF received amplitudes and mesospheric temperatures. Furthermore, the possible causalities of high negative correlation between mesospheric temperature and VLF amplitude are first, the atmosphere in a major sense gets heated from the bottom (from the ground surface) and direct heating from the sun is rather low relative to surface heating. Since mesosphere is far away from the surface, the heating is rather low. Second, mesosphere has a lot of radiative cooling. The CO_2 gas present here emits a lot of thermal radiation into space. The energy for this radiative release is obtained by the energy transfer from other molecules at lower levels. Due to lesser density, the radiative cooling is large enough to cause a steep decrease in temperature. Since the stratospheric temperature has positive correlation with VLF amplitude in the mid-latitude path as mentioned by *Pal and Hobara*, [2016]. Thus, the heat propagation from the stratosphere to mesosphere regions has opposite characteristic due to radiative cooling in mesosphere layer.

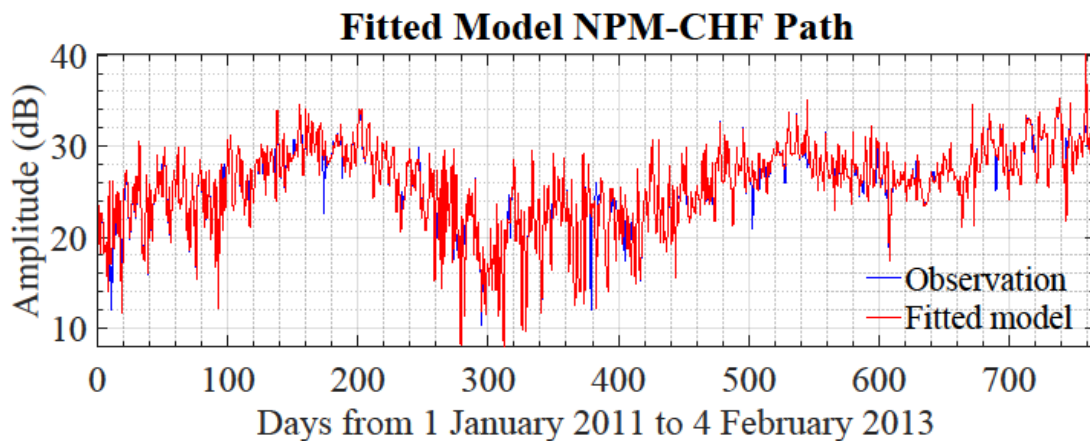


Figure 5.21: The fitted model predictions of NARX NN model of daily nighttime mean amplitude of VLF waves mid-latitude path with three-day of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 1 January 2011 to 4 February 2013 (VLF observation-blue solid; The fitted model-red dotted).

The NARXNN model with three days of input-memory and two hundred neurons in the hidden layer using LMANN algorithm is used to predict the daily nighttime of VLF electric field amplitude mid-latitude path. The fitted model (inside the training period) with the time interval from 1 January 2011 to 4 February 2013 show in red curve and observation in blue curve as depicted in Figure 5.21. As a result, the fitted model has a good agreement with the

original data for each path. The constructed model worked well for prediction as represented by Pearson correlation coefficient (r) is 0.951.

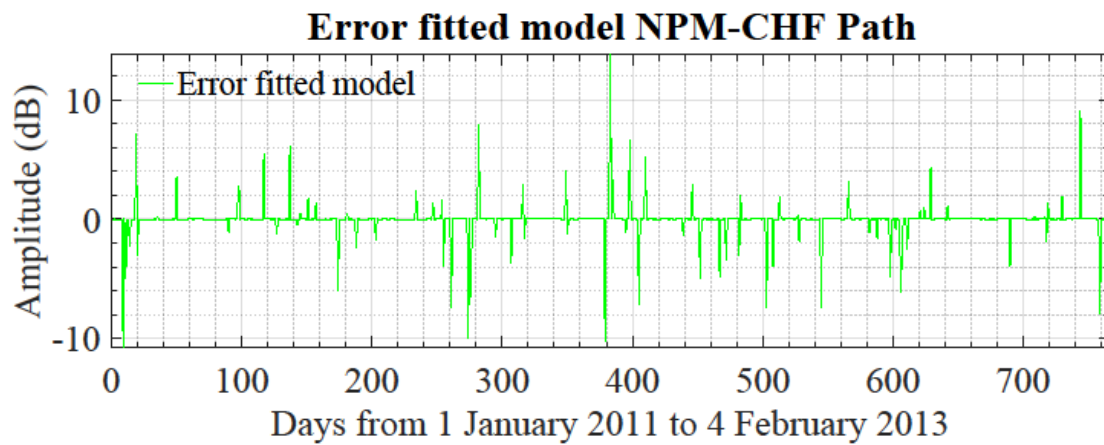


Figure 5.22: Error fitted model predictions of NARX NN model of daily nighttime mean amplitude of VLF waves mid-latitude path over the time interval from 1 January 2011 to 4 February 2013.

Furthermore, the prediction error for 766 days data point is shown graphically in Figure 5.22. The error varies from -10.7 dB to 13.7 dB and the RMSE is 1.45 dB. We have to consider a few dates with a relatively large error which mean big discrepancy between observed and fitted model values in low-mid-latitude Chofu station. This may be due to the physical factors other than we consider in the model inputs. Therefore, these observed values are considered to be anomalies due to unknown physical reason in the proposed model.

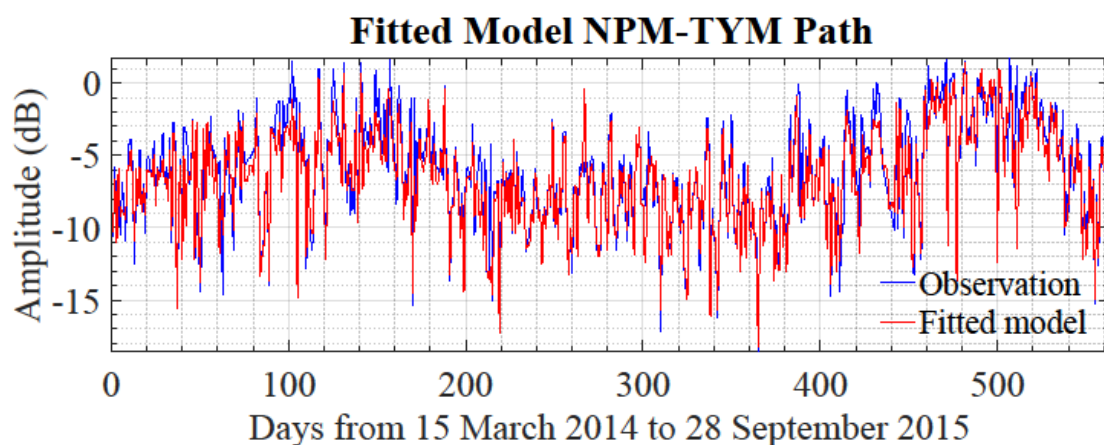


Figure 5.23: The fitted model predictions of NARX NN model of daily nighttime mean amplitude of VLF waves mid-latitude path with three-day of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 15 March 2014 to 28 September 2015 (VLF observation-blue solid; The fitted model-red dotted).

As described in Section 6.2.2.1, for comparative study and to examine the capability of our NARXNN model, different datasets are used. The results are illustrated in Figure 5.23. The blue curve denotes the actual VLF amplitude values and the red curve shows the predicted ones. The fitted model also has a good agreement as Chofu station with the original data for the low-mid-latitude path. The NARXNN predictor model successful for prediction as represented by Pearson correlation coefficient (r) of 0.922.

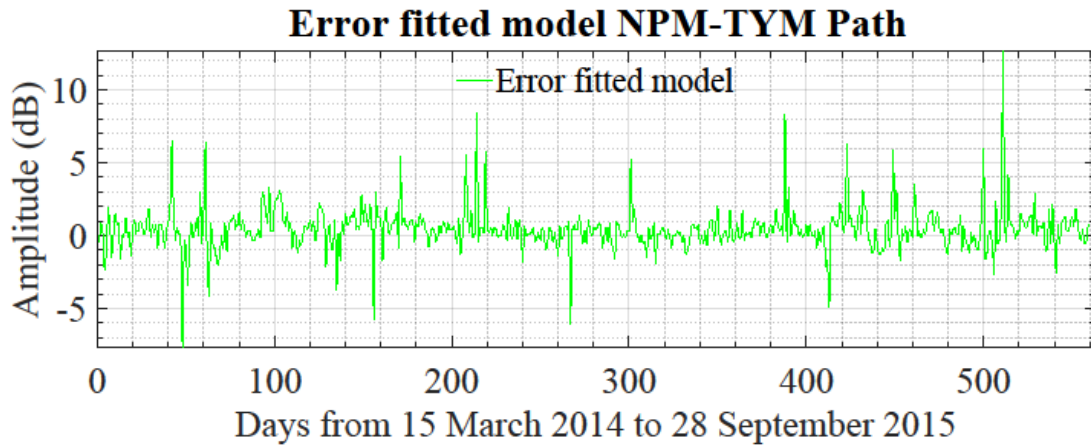


Figure 5.24: Error fitted model predictions of NARX NN model of daily nighttime mean amplitude of VLF waves mid-latitude path over the time interval from 15 March 2014 to 28 September 2015.

Moreover, the prediction error for 563 days data point is shown graphically in Figure 5.24. The error varies from -7.6 dB to 12.6 dB and the RMSE is 1.56 dB. We have to consider a few dates with a relatively large error which mean big discrepancy between observed and fitted model values in mid-latitude Tsuyama station. This may be due to the physical factors other than we consider in the model inputs. Therefore, these observed values are considered to be anomalies due to unknown physical reason in the proposed model.

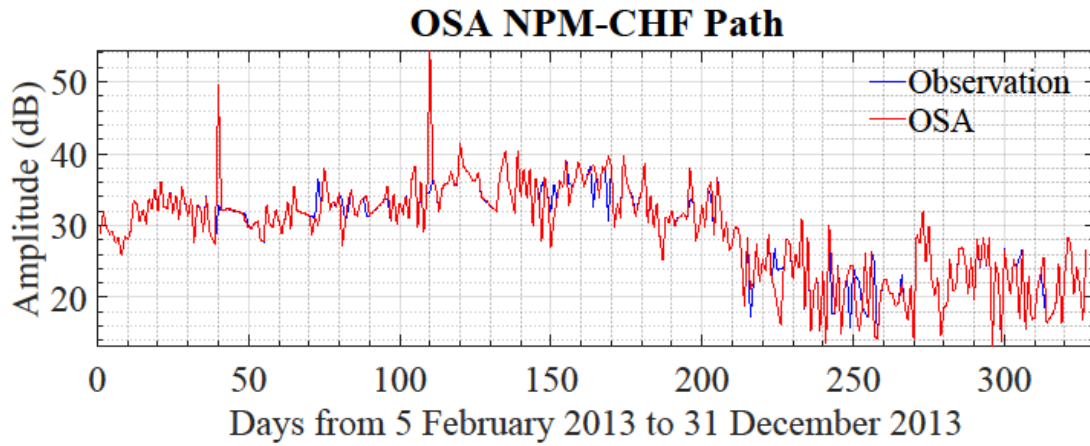


Figure 5.25: One Step (1 day) Ahead (OSA) predictions of NARX NN model of daily nighttime mean amplitude of VLF waves mid-latitude path with three-day of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 1 January 2011 to 4 February 2013 (VLF observation-blue solid; The fitted model-red dotted).

Figure 5.25 shows the OSA prediction with the data set outside the training period. The correlation coefficient for the prediction remains high value as represented by r of 0.940.

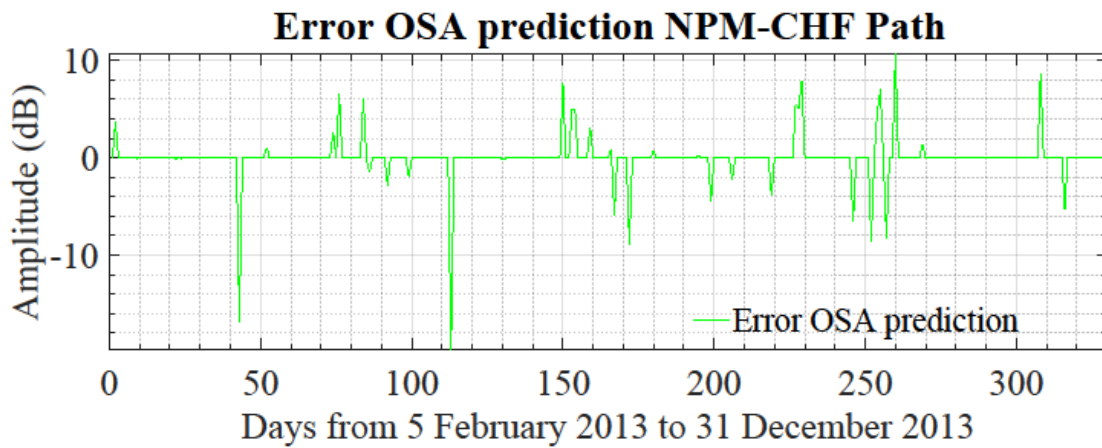


Figure 5.26: Error One Step (1 day) Ahead (OSA) predictions of NARX NN model of daily nighttime mean amplitude of VLF waves mid-latitude path over the time interval from 1 January 2011 to 4 February 2013.

Moreover, the prediction error for 330 days data point outside the training period is shown graphically in Figure 5.26. The error varies from -19.6 dB to 10.5 dB and the RMSE is 2.22 dB.

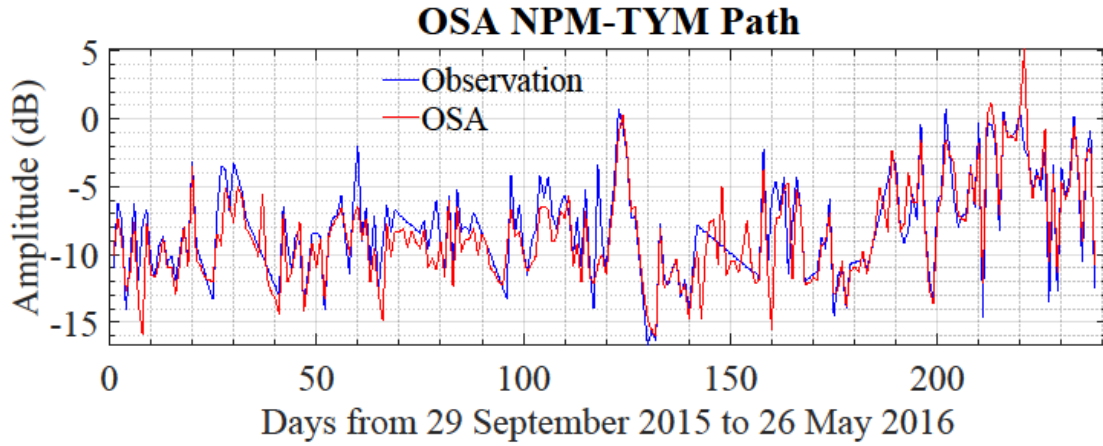


Figure 5.27: One Step (1 day) Ahead (OSA) predictions of NARX NN model of daily nighttime mean amplitude of VLF waves mid-latitude path with three-day of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 29 September 2015 to 26 May 2016 (VLF observation-blue solid; Prediction-red dotted).

Further, to examine the capability of our NARXNN model, we feed the built model with different datasets from different receiving station over the time interval between 29 September 2015 to 26 May 2016. We still use NARXNN model equation for mid-latitude path Tsuyama station with three days delay time $\{d_u = 3, d_y = 3\}$. The results are illustrated in Figure 5.27. The blue curve denotes the actual VLF amplitude values and the red curve shows the predicted ones. The OSA still has a good agreement as a fitted model with the original data for low-mid-latitude path. The NARXNN predictor model successful for prediction outside the training period as shown by Pearson correlation coefficient (r) of 0.870.

Moreover, the prediction error of OSA outside the training period for 241 days data point is shown graphically in Figure 5.28. The error varies from -7.3 dB to 9.7 dB and the RMSE is 1.90 dB. This result has a relatively good in average compared with Chofu station. This condition may be due to the physical factors other than we consider in the model inputs and also depend on the signals interferences on receiving site

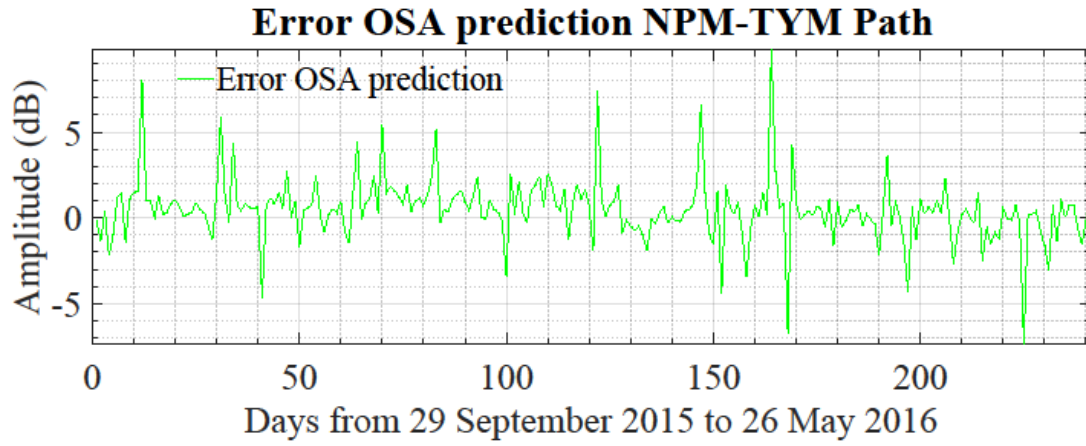


Figure 5.28: Error One Step (1 day) Ahead (OSA) predictions of NARX NN model of daily nighttime mean amplitude of VLF waves mid-latitude path over the time interval from 29 September 2015 to 26 May 2016.

5.2.2.3 High-latitude VLF propagation path

$$\begin{aligned}
 VLF(k) = F [& 2.108VLF(k-1) - 1.178VLF(k-2) \\
 & + 1.566VLF(k-3) - 1.390ST(k-1) \\
 & + 0.982ST(k-2) + 0.996ST(k-3) \\
 & - 2.857C(k-1) - 2.604C(k-2) \\
 & + 5.076C(k-3) - 3.756TCO(k-1) \\
 & + 0.183TCO(k-2) + 0.263TCO(k-3) \\
 & + 5.399Dst(k-1) - 2.419Dst(k-2) \\
 & + 2.025Dst(k-3) - 4.081AE(k-1) \\
 & + 1.105AE(k-2) + 1.133AE(k-3) \\
 & - 2.797Kp(k-1) - 0.034Kp(k-2) \\
 & + 2.687Kp(k-3) + 1.130MT(k-1) \\
 & + 1.043MT(k-2) + 2.459MT(k-3) \\
 & + 0.592F10.7(k-1) + 1.877F10.7(k-2) \\
 & - 0.020F10.7(k-3) + 1.069]
 \end{aligned} \tag{5.14}$$

As represented in equation (5.14) that NLK-CHF path has the first significant factor which represents with the Dst index 1 day before the given day $\{Dst(k-1)\}$ as indicated by the coefficient of 5.399. The second influential factor is the cosmic ray with 3 days before the given day $\{C(k-3)\}$ with a weighting coefficient of 5.076. The AE index of 1 day before the given day $\{AE(k-1)\}$ becomes the third influential factor as represented by the coefficient

of 4.081. Finally, the total column ozone with 1 day before the given day $\{TCO(k-1)\}$ is the fourth significant parameter with the coefficient of 3.756.

Moreover, the TYM station NARXNN model equation can be represented in equation (5.15) and for the first four significant terms for NWC-TYM path are described here. The first influential factor is the Dst index with one day before the given day $\{Dst(k-1)\}$ as represented by the coefficient of 3.992. The cosmic ray with one day before the given day $\{C(k-1)\}$ becomes the second significant factor with the coefficient of 3.396. The third influential factor is the AE index with one day before the given days $\{AE(k-1)\}$ with a weighting coefficient of 2.512. The last one is total column ozone with one day before the given day $\{TCO(k-1)\}$ as indicated by the coefficient of 2.354.

$$\begin{aligned}
 VLF(k) = F [& 1.737VLF(k-1) + 0.797VLF(k-2) \\
 & + 0.284VLF(k-3) - 0.450ST(k-1) \\
 & + 0.964ST(k-2) + 0.840ST(k-3) \\
 & + 3.396C(k-1) - 1.764C(k-2) \\
 & - 2.120C(k-3) + 0.301TCO(k-1) \\
 & + 2.354TCO(k-2) + 0.292TCO(k-3) \\
 & + 3.992Dst(k-1) + 1.835Dst(k-2) \\
 & + 1.274Dst(k-3) + 2.512AE(k-1) \\
 & + 1.590AE(k-2) - 1.100AE(k-3) \\
 & - 2.056Kp(k-1) + 0.647Kp(k-2) \\
 & - 0.623Kp(k-3) - 1.223MT(k-1) \\
 & + 0.178MT(k-2) - 0.131MT(k-3) \\
 & - 1.350F10.7(k-1) - 0.695F10.7(k-2) \\
 & - 0.048F10.7(k-3) + 0.225]
 \end{aligned} \tag{5.15}$$

Figure 5.29 and Figure 5.30 show the four of most relative significant factor which has influences in the model predictor high-latitude path. For the most relative significant parameter, both the receiving station between Chofu and Tsuyama have the same variable of Dst index one day before the given day. Tsuyama station has a bigger percentage compare with Chofu station and it shows with the value of 10.43% and 11.55% for NLK-CHF and NLK-TYM respectively. Further, the second relative significant factor, for Chofu station is cosmic ray with three days before the given day as represented in Figure 5.29 with the value of 9.81%

and Tsuyama station is cosmic ray with one day before the given day as illustrated in Figure 5.30 with the value of 9.83%. This difference in the delay time in the low-mid-latitude path may be due to unknown parameter outside the considered inputs. Moreover, the third relative significant factor has the same parameter in both receiving stations with AE index one day before the given day and the value of 7.88% and 7.27% as depicted in Figure 5.29 and Figure 5.30 respectively. NLK-CHF has a higher percentage compared to NLK-TYM with the different value of 0.61%. Finally, the fourth relative significant parameter also different between NLK-CHF and NWC-TYM paths as shown total column ozone one day before the given day with value of 7.26% and two days before the given day with value of 6.81% respectively

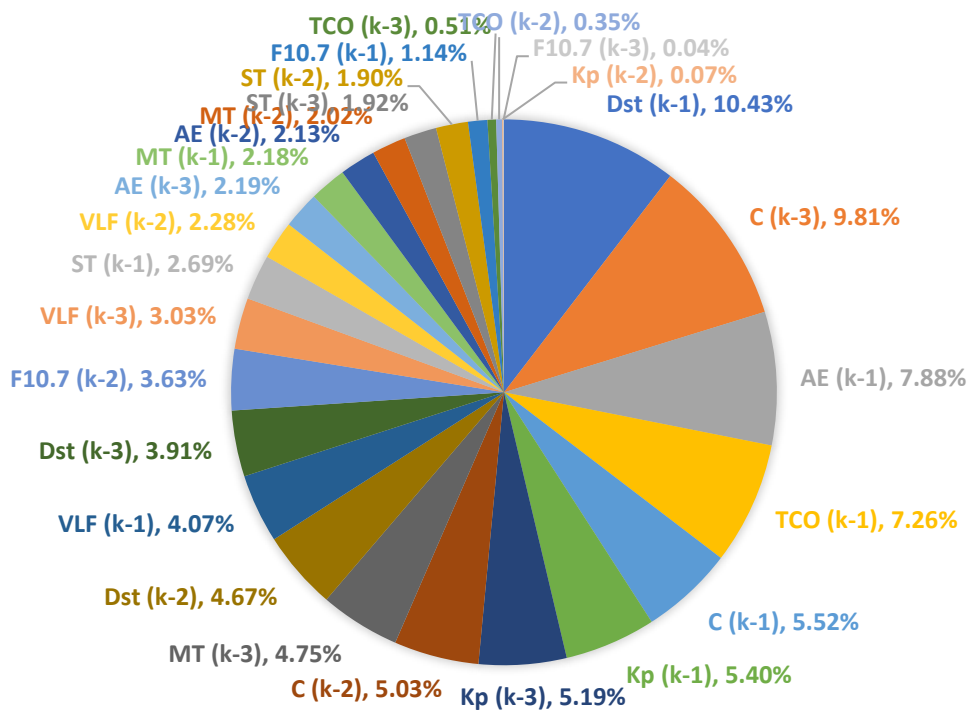


Figure 5.29: The relative significant parameter in high-latitude NARXNN model of the daily nighttime mean values of VLF electric field amplitude Chofu station.

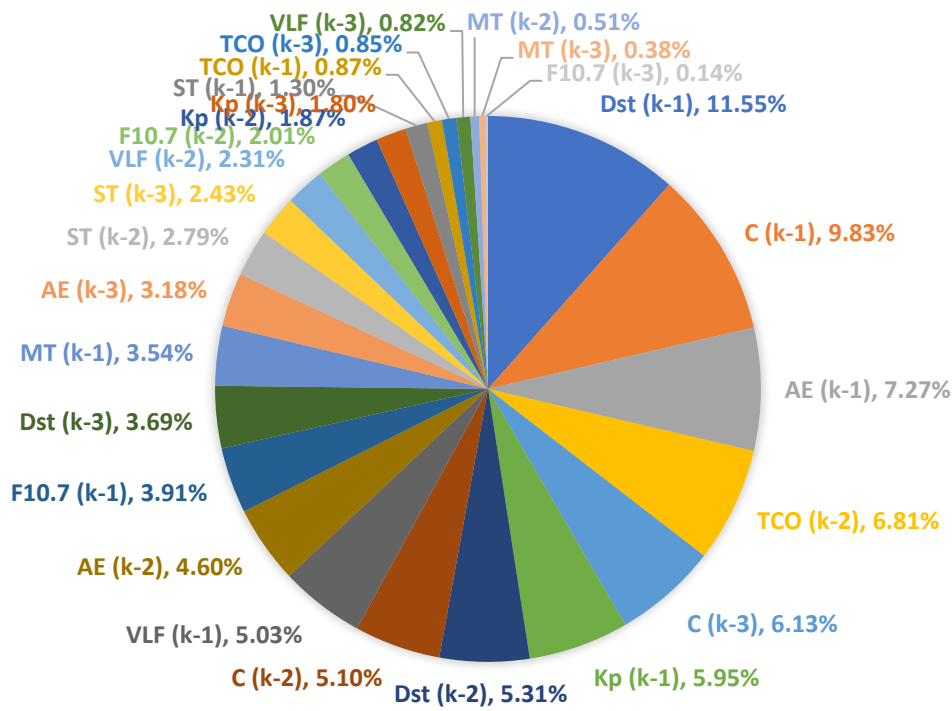


Figure 5.30: The relative significant parameter in high-latitude NARXNN model of the daily nighttime mean values of VLF electric field amplitude Tsuyama station.

Dst index is the most significant variable in NLK-CHF path with the large coefficients that contributed in the prediction model. *Tatsuta et al.* [2015] mentioned that in the high-latitude path Dst index has a high correlation with VLF amplitude anomaly. The deviation of the daily mean nighttime of VLF amplitude from the corresponding mean values from the past 15 days has been used as a daily dependence of the VLF amplitude [*Tatsuta et al.*, 2015]. Precipitations of energetic electrons due to the pitch-angle scattering and diffusion due to the cyclotron resonance at the lower edge of the inner radiation belt may disturb D/E region ionosphere during the storm time period [*Jain and Singh*, 1990]. The VLF amplitude depression and fluctuation during the geomagnetic storm period correspond to the electron density enhancement in D-region caused by high-energy auroral electron precipitation [*Cummer et al.*, 1997] which leads the change in VLF trend and nighttime fluctuation values. Further VLF perturbation affected by the magnetic disturbance approximately within 1 day [*Abdu et al.*, 1981].

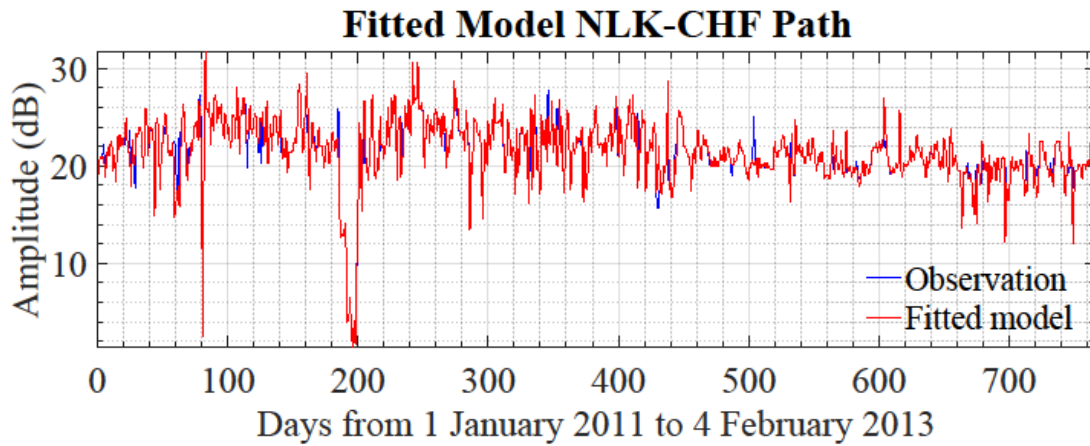


Figure 5.31: The fitted model predictions of NARX NN model of daily nighttime mean amplitude of VLF waves high-latitude path with three-day of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 1 January 2011 to 4 February 2013 (VLF observation-blue solid; The fitted model-red dotted).

Cosmic rays flux depends in the solar activity and the geomagnetic field quantities enhanced ionization in the lower ionosphere in the high-latitude regions. Moreover, the solar cosmic rays appear to be essentially isotropic in the surrounding of the Earth and remain in this region for several days and provide ionization over this period [Webber, 1962]. The ionization produced at nighttime by cosmic rays forms a layer near 95 km which is capable of reflecting VLF waves [Moler, 1960]. The possible mechanism is that an ion produced by cosmic rays ionized the positive ion of N_2^+ below lower ionosphere whereas above this region O_2^+ and NO^+ become the dominant positive ions, thus the average effective electron-ion recombination coefficient gradually changes from $\sim 3 \times 10^{-7}$ cm³/sec below 65 km to $\sim 3 \times 10^{-8}$ cm³/sec near the 100 km [Nicolet and Aikin, 1960]. Since the ionization by cosmic ray is constant, three-day time delay may be due to unknown external forcing from atmosphere or below D-region.

AE index also contributes to the high-latitude path prediction model. Similar results have been obtained by Tatsuta *et al.*, [2015] that AE index in the high-latitude path has a high correlation with VLF amplitude anomaly. Furthermore, total column ozone is recognized as one of the significant factors contributing to VLF amplitude prediction. The correlation between D-region electron density and the corresponding variability in the total ozone is high over the mid-latitude [Abdu and Angreji, 1974]. Considering the relative position of NLK-CHF path within the high-middle-latitudes region, there is a possibility that the stratospheric ozone participates in the VLF perturbation process.

The NARXNN model with 3 days of input-memory and 200 neurons in the hidden layer using LMANN algorithm is used to predict the daily nighttime of VLF electric field

amplitude for the high-latitude path. The fitted model (inside the training period) with the time interval from 1 January 2011 to 4 February 2013 show in red curve and observation in blue curve as shown in Figure 5.31. As a result, the fitted model has a good agreement with the original data for each path. The built model performed well for prediction as represented by Pearson correlation coefficient (r) is 0.936.

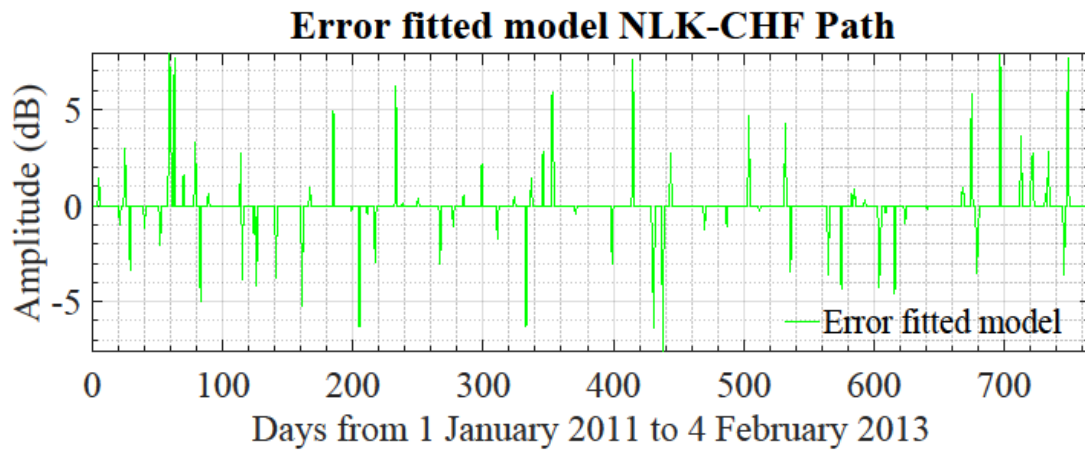


Figure 5.32: Error fitted model predictions of NARX NN model of daily nighttime mean amplitude of VLF waves high-latitude path over the time interval from 1 January 2011 to 4 February 2013.

Furthermore, the prediction error for 766 days data point is shown graphically in Figure 5.32. The error varies from -7.5 dB to 7.8 dB and the RMSE is 1.18 dB. We have to consider a few dates with a relatively large error which mean big discrepancy between observed and fitted model values in low-mid-latitude Chofu station. This may be due to the physical factors other than we consider in the model inputs. Therefore, these observed values are considered to be anomalies due to unknown physical reason in the proposed model.

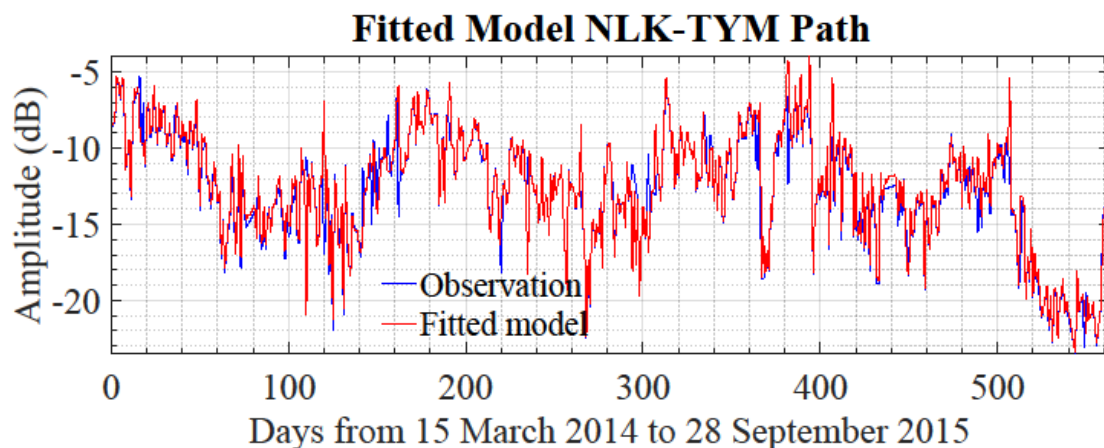


Figure 5.33: The fitted model predictions of NARX NN model of daily nighttime mean amplitude of VLF waves high-latitude path with three-day of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 15 March 2014 to 28 September 2015 (VLF observation-blue solid; The fitted model-red dotted).

As described in the section before, for comparative study and to examine the capability of our NARXNN model, different datasets are used. The results are illustrated in Figure 5.33. The blue curve denotes the actual VLF amplitude values and the red curve shows the predicted ones. The fitted model also has a good agreement with Chofu station with the original data for the low-mid-latitude path. The NARXNN predictor model successful for prediction as represented by Pearson correlation coefficient (r) of 0.938.

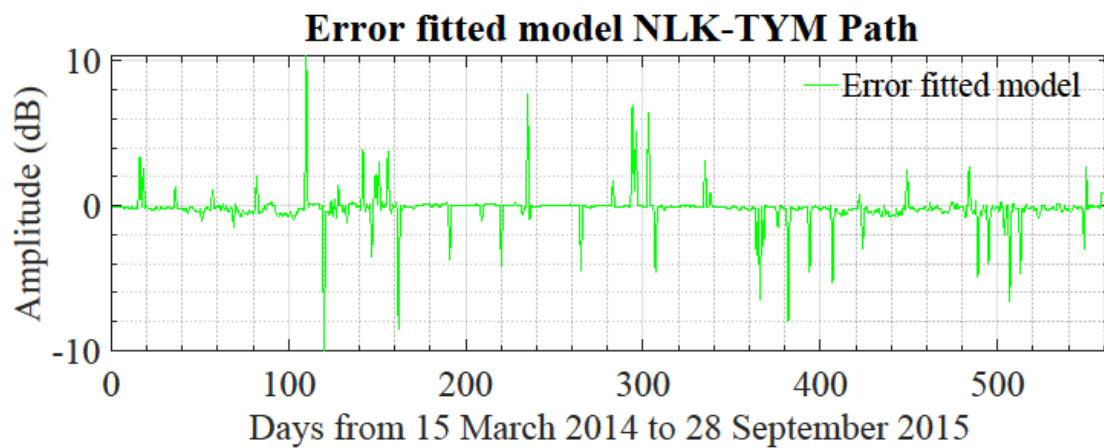


Figure 5.34: Error fitted model predictions of NARX NN model of daily nighttime mean amplitude of VLF waves high-latitude path over the time interval from 15 March 2014 to 28 September 2015.

Further, the prediction error for 563 days data point is shown graphically in Figure 5.34. The error varies from -10 dB to 10.3 dB and the RMSE is 1.32 dB. We have to consider a few dates with a relatively large error which mean big discrepancy between observed and fitted model values in mid-latitude Tsuyama station. This may be due to the physical factors other than we consider in the model inputs. Therefore, these observed values are considered to be anomalies due to unknown physical reason in the proposed model.

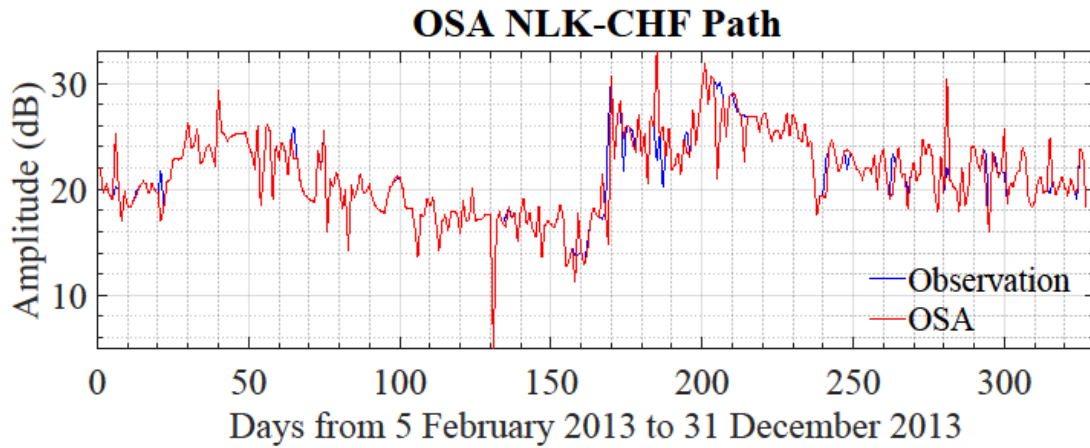


Figure 5.35: One Step (1 day) Ahead (OSA) predictions of NARX NN model of daily nighttime mean amplitude of VLF waves high-latitude path with three-day of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 5 February 2013 to 31 December 2013 (VLF observation-blue solid; Prediction-red dotted).

Figure 5.35 shows the OSA prediction with the data set outside the training period. The correlation coefficient for the prediction remains high value as represented by r of 0.931.

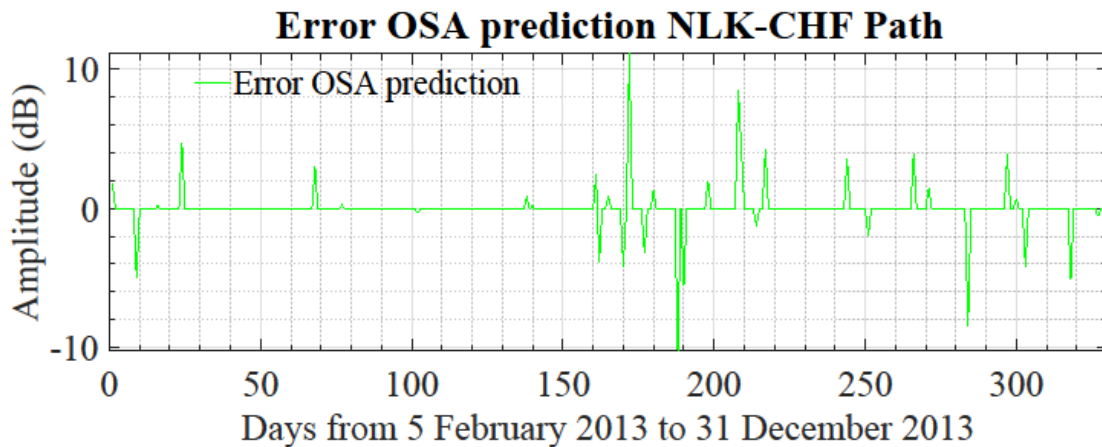


Figure 5.36: Error One Step (1 day) Ahead (OSA) predictions of NARX NN model of daily nighttime mean amplitude of VLF waves high-latitude path over the time interval from 1 January 2011 to 4 February 2013.

Moreover, the prediction error for 330 days data point outside the training period is shown graphically in Figure 5.36. The error varies from -10.2 dB to 11 dB and the RMSE is 1.34 dB.

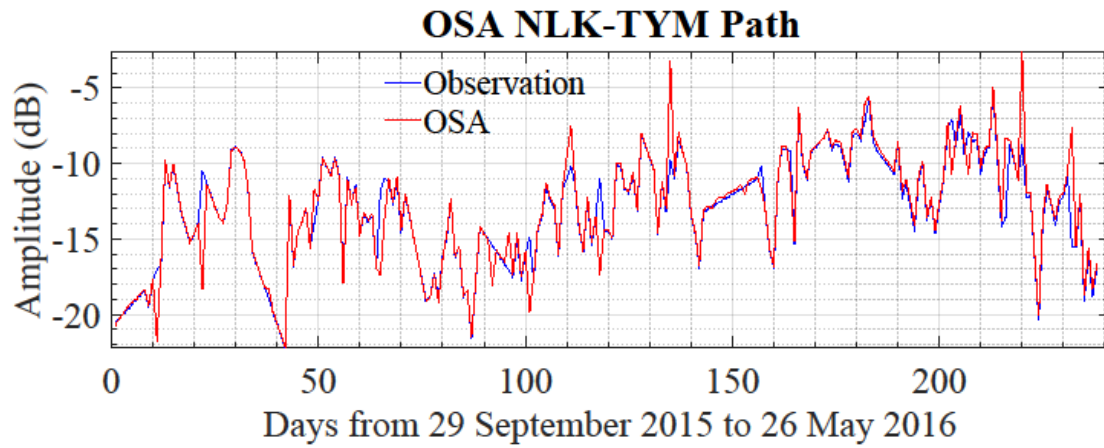


Figure 5.37: One Step (1 day) Ahead (OSA) predictions of NARX NN model of daily nighttime mean amplitude of VLF waves high-latitude path with three-day of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 29 September 2015 to 26 May 2016 (VLF observation-blue solid; Prediction-red dotted).

Further, to examine the capability of our NARXNN model, we feed the built model with different datasets from different receiving station over the time interval between 29 September 2015 to 26 May 2016. We still use NARXNN model equation for high-latitude path Tsuyama station with three days delay time $\{d_u = 3, d_y = 3\}$. The results are illustrated in Figure 5.37. The blue curve denotes the actual VLF amplitude values and the red curve shows the predicted ones. The OSA still has a good agreement as a fitted model with the original data for low-mid-latitude path. The NARXNN predictor model successful for prediction outside the training period as shown by Pearson correlation coefficient (r) of 0.933.

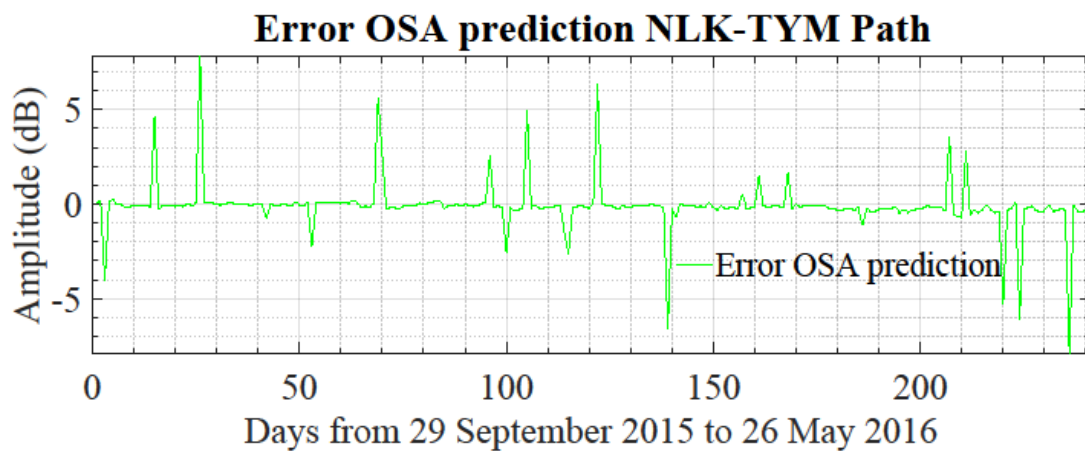


Figure 5.38: Error One Step (1 day) Ahead (OSA) predictions of NARX NN model of daily nighttime mean amplitude of VLF waves high-latitude path over the time interval from 29 September 2015 to 26 May 2016.

Moreover, the prediction error of OSA outside the training period for 241 days data point is shown graphically in Figure 5.38. The error varies from -7.9 dB to 7.7 dB and the RMSE is 1.35 dB. This result has a relatively good in average compared with Chofu station. This condition may be due to the physical factors other than we consider in the model inputs and also depend on the signals interferences on receiving site.

Table 5.2: Fit Pearson's correlation coefficient (r) and root mean squared error (RMSE) of daily VLF prediction both Chofu and Tsuyama stations.

Path	Chofu				Tsuyama			
	Fitted model		OSA		Fitted model		OSA	
	r	RMSE	r	RMSE	r	RMSE	r	RMSE
Low-mid-latitude	0.913	1.50 dB	0.910	1.68 dB	0.915	0.98 dB	0.909	1.10 dB
Mid-latitude	0.951	1.45 dB	0.940	2.22 dB	0.922	1.56 dB	0.870	1.90 dB
High-latitude	0.936	1.18 dB	0.931	1.34 dB	0.938	1.32 dB	0.933	1.35 dB

Table 5.2 summarize the performance of Chofu and Tsuyama VLF prediction model inside and outside the training periods. NARXNN prediction model using LMANN algorithm with the input memory of 3 days before the given day and 200 neurons in the hidden layer is found to have a good performance for different latitude paths. As shown in Table 5.2, Pearson's correlation coefficient both receiver station for different latitude greater than 0.87 and RMSE less than 2.25 dB.

5.3 Multi-step Ahead Prediction (MSA)

In this thesis, based on the results in Section 6.2.2 which the OSA NARXNN nonlinear model for low-, mid-, and high-latitude paths to develop a multi-step ahead (MSA) model to predict the daily nighttime mean value of VLF electric field amplitude. We evaluate our MSA model using two different datasets, first for Chofu station with the datasets range from 21 August 2013 to 18 December 2013 (outside of the training period) and second for Tsuyama station with the datasets over the time interval between 26 January 2016 and 24 May 2016 (outside of the training period). For time series prediction strategy, we use hybrid NARXNN

with a combination of recursive and direct strategies (DirRec strategy). In DirRec strategy, for each future value, the constructed model needs to be established separately. Further, the initial prediction formula is executed to obtain the initial estimation value, then the value as for be the input of next step. In addition, this process will continue until reaching the prediction horizon.

The NARXNN DirRec strategy constructed model was used to predict five and ten days ahead of the daily nighttime mean value of VLF electric field amplitude for three different latitude paths namely low-mid-latitude path (NWC-CHF and NWC-TYM), middle-latitude path (NPM-CHF and NPM-TYM), and high-latitude path (NLK-CHF and NLK-TYM).

5.3.1 Architecture of NARXNN MSA Modeling

A multi-step ahead VLF electric field prediction can be described as an estimation of future VLF electric field value consists of predicting the next H values $[y_{N+1}, \dots, y_{N+H}]$ of a historical time series $[y_1, \dots, y_N]$ composed of N observations, where $H > 1$ denotes the prediction horizon. We used Direct Recursive (DirRec) strategy [Sorjamaa and Lendasse, 2006] for multi-step ahead prediction. DirRec computes the predictions with different models for every time step, it enlarges the set of inputs by adding variable corresponding to the predictions of the previous step. In other term, the DirRec strategy learns H models f_h from the time series $[y_1, \dots, y_N]$ where

$$y_{k+h} = f_h(y_{k+h-1}, \dots, y_{k-d+1}) \quad (5.16)$$

with $k \in \{d, \dots, N - H\}$, f_h NARXNN model and $h \in \{1, \dots, H\}$.

To obtain the multi-step ahead prediction, the H learned models are used as follows:

$$\hat{y}_{N+h} \begin{cases} \hat{f}_h(y_N, \dots, y_{N-d+1}) & \text{if } h = 1 \\ \hat{f}_h(\hat{y}_{N+h-1}, \dots, \hat{y}_{N+1}, y_N, \dots, y_{N-d+1}) & \text{if } h \in \{2, \dots, H\} \end{cases} \quad (5.17)$$

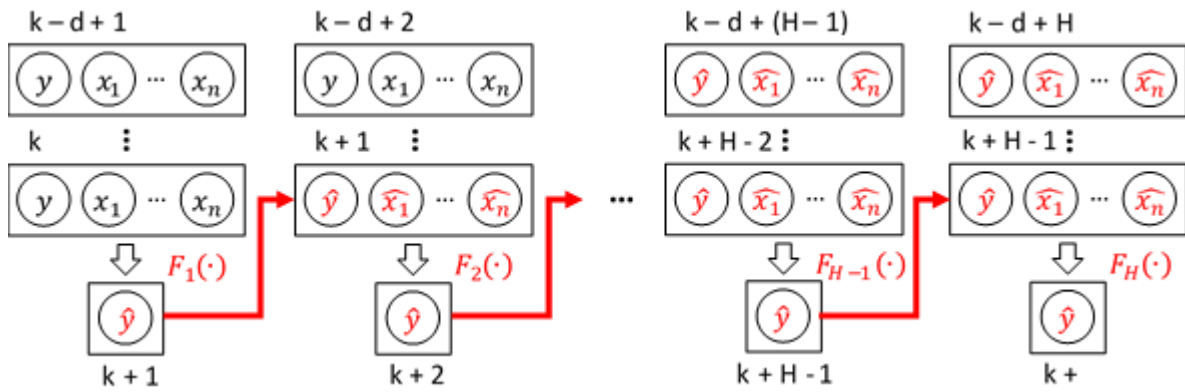


Figure 5.39: DirRec strategy for multi-step ahead prediction.

The NARXNN constructed a model is trained to perform a one-step ahead prediction and when predicting n steps ahead, the first step by applying the NARXNN OSA model. Subsequently, the independence predicted value for each input are used as part of the new input variables for predicting the next step. The input data set increase with one more input every time step. The complexity of the model increases linearly with more inputs and prediction error are fed into the NARXNN model. In this thesis, the constructed model based on the NARXNN one-step ahead prediction with 3 delay time and 200 neurons using LMANN algorithm devised a NARXNN multi-step ahead prediction. The NARXNN MSA prediction architecture build based on the DirRec strategy with m external inputs and 1 output represented diagrammatically as Figure 4. Let $u(k)$ denote the $m \times d$ input vector at discrete time k and $u(k+n)$ denote the $m \times d$ input vector n -step later at time $k+n$. For $y(k)$ denote the corresponding $1 \times d$ vector at discrete time k and $y(k+n)$ denote the corresponding $1 \times d$ vector one-step later at time $k+n$.

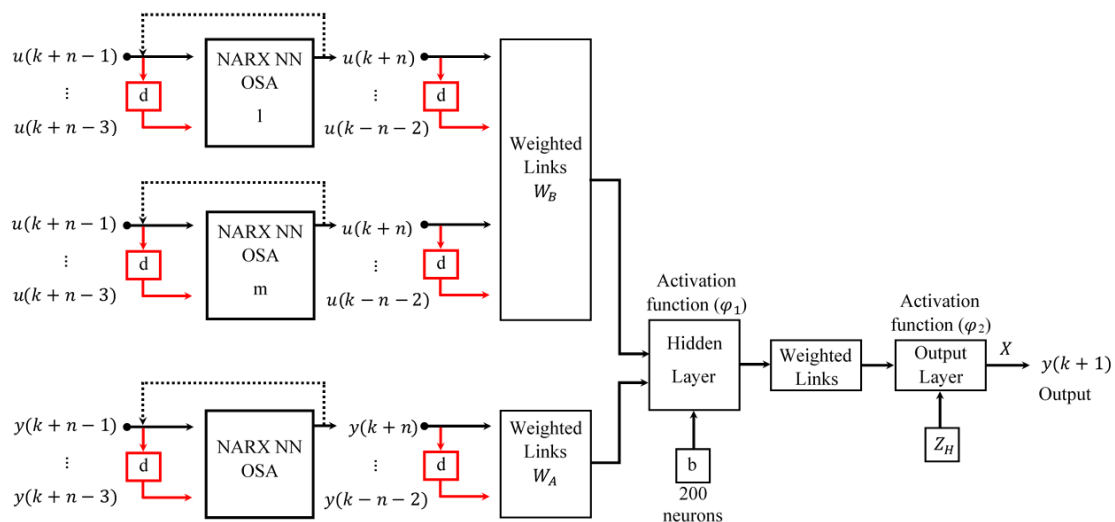


Figure 5.40: NARXNN MSA prediction architecture.

5.3.2 Prediction Results

After implementing NARXNN constructed model for OSA prediction based on data set inside and outside training period, we demonstrate MSA prediction. The same model with 3 days of input-memory and 200 neurons in the hidden layer using LMANN algorithm is used to initial structure of MSA prediction model. Since the DirRec strategy is applied In this thesis, the prediction for the next step will compute new model structure. The predicted value use as a part of the new input for training the NARXNN and predicting the next step. We continue this manner until the entire horizon is predicted. Furthermore, the build model is used for performing five and ten days ahead prediction with 120 days data sets from 21 August 2013 to 18 December 2013 (outside training period) for Chofu station and 26 January 2016 to 24 May 2016 (outside training period) for Tsuyama station.

In this section, we first demonstrate MSA prediction for different latitude path of VLF wave propagation and also two different receiving station for a comparative study. We will compare the actual and the predicted values which are the NARXNN outputs. We also study them statistically and depict the errors.

5.3.2.1 Low-Mid-Latitude VLF propagation path

To represent the predicting accuracy of NARXNN MSA prediction models in low-mid-latitude VLF propagation path, the curve plot of the observed data (blue curve) versus the predicted value of 5 days ahead in the 120 days data sets are shown in Figure 5.41 for NWC-CHF path. As a result, almost for all VLF paths of the predicted value suitably follow the observed value as indicated by a correlation coefficient is 0.79.

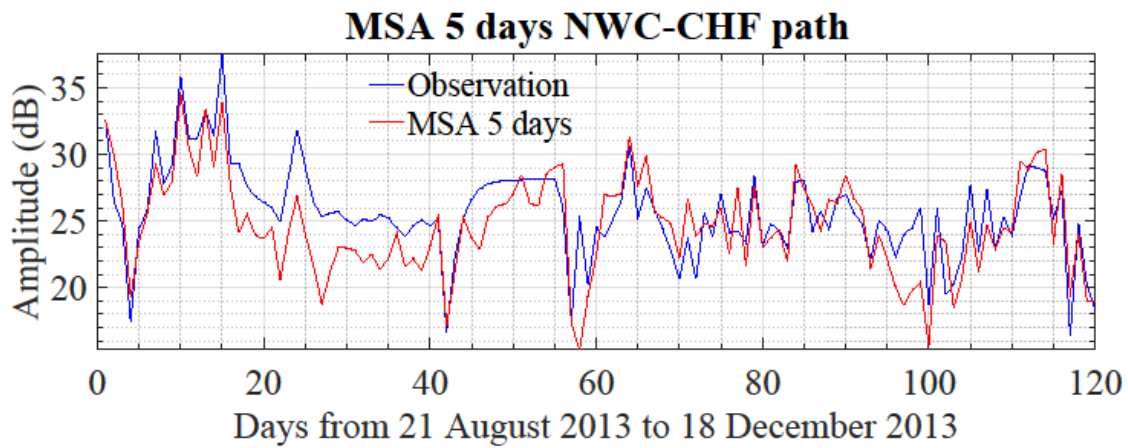


Figure 5.41: Multi Step (5 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves low-mid-latitude path with three days of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 21 August 2013 to 18 December 2013 (VLF observation-blue solid; Prediction-red dotted).

Further, the graphical representation of the error MSA 5 days prediction for NWC-CHF path (low-mid-latitude) in Figure 5.42. Further, the statistical performance calculates to show the capability of the model predictor. The minimum value of the prediction error is -3.90 dB, the maximum value is 9.94 dB, the mean value is 1.00 dB, the standard deviation is 2.37 dB the error prediction shows in RMSE of 2.60 dB (NWC-CHF). The proposed model shows the ability to predict 5 days ahead as indicate with the correlation coefficient is reasonably good and the RMSE also relatively small.

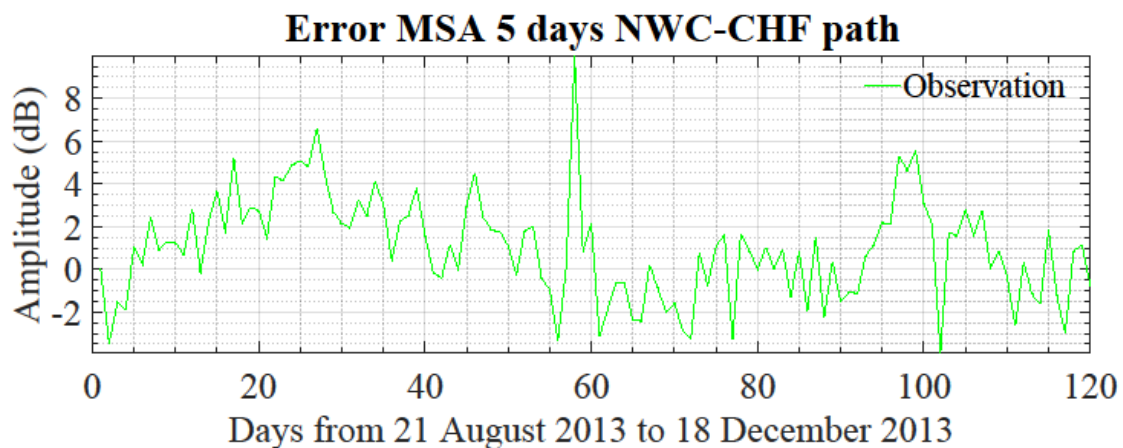


Figure 5.42: Error Multi Step (5 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves low-mid-latitude path over the time interval from 21 August 2013 to 18 December 2013.

Moreover, to examine the capability of our NARXNN DirRec model predictor, we use different datasets from different receiving station over the time interval between 15 March 2014 and 28 September 2015. The results are depicted in Figure 5.43. The blue curve denotes the actual VLF amplitude values and the red curve shows the predicted value. The constructed model successful for representing dynamic changing of VLF amplitude even for 5 days ahead as represented by Pearson correlation coefficient (r) of 0.801.

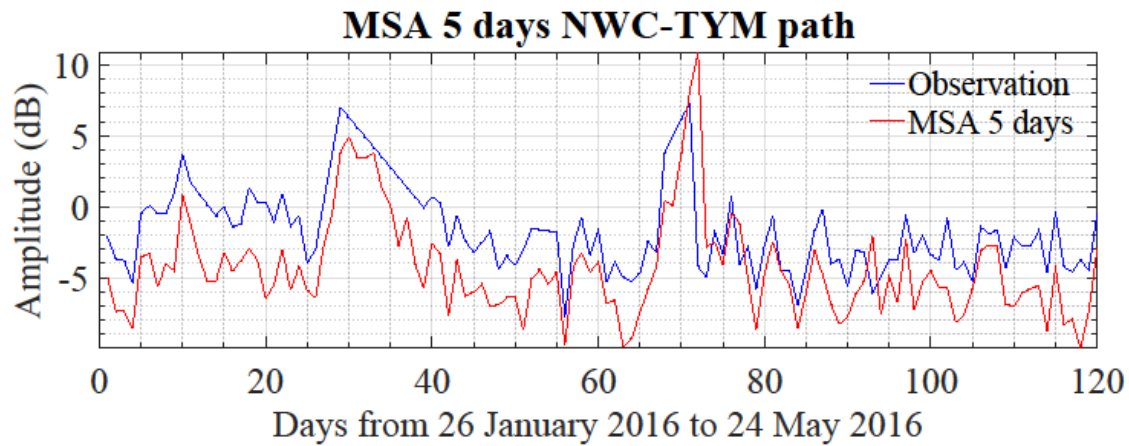


Figure 5.43: Multi Step (5 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves low-mid-latitude path with three days of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 26 January 2016 to 24 May 2016 (VLF observation-blue solid; Prediction-red dotted).

Further, the graphical representation of the error MSA 5 days prediction for NWC-TYM path (low-mid-latitude) in Figure 5.44. Further, the statistical performance calculates to show the capability of the model predictor. The minimum value of the prediction error is -14.97 dB, the maximum value is 6.75 dB, the mean value is 2.72 dB, the standard deviation is 2.33 dB the error prediction shows in RMSE of 3.57 dB (NWC-TYM). The proposed model shows the ability to predict 5 days ahead as indicate with the correlation coefficient is reasonably good and the RMSE also relatively small.

The same data sets are used to perform the 10 days ahead prediction of the daily nighttime of VLF electric field amplitude. Figure 5.45 shows the 10 days ahead prediction for low-mid-latitude (NWC-CHF). The blue curve denotes the actual VLF amplitude values and the red curve shows the predicted value. The prediction performances for 10 days ahead are represented by a correlation coefficient of 0.63.

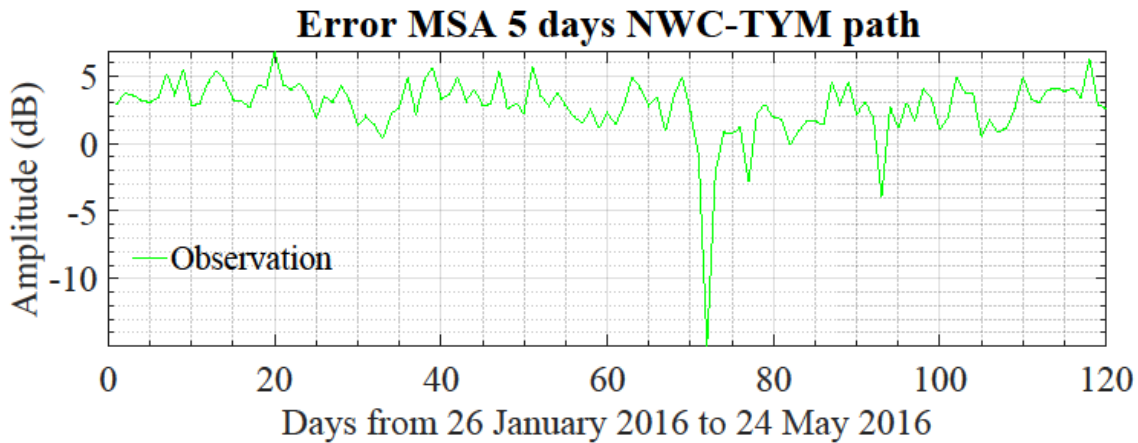


Figure 5.44: Error Multi Step (5 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves low-mid-latitude path over the time interval from 26 January 2016 to 24 May 2016.

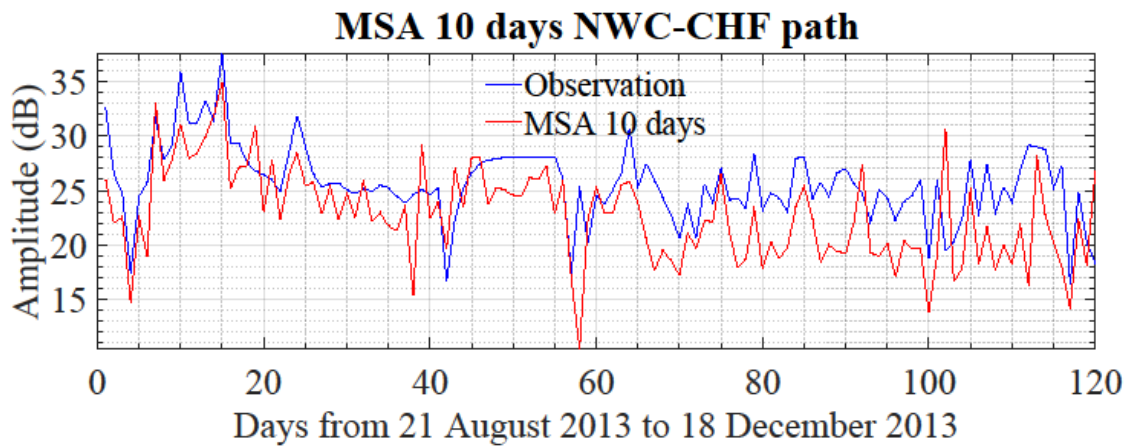


Figure 5.45: Multi Step (10 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves low-mid-latitude path with three days of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 21 August 2013 to 18 December 2013 (VLF observation-blue solid; Prediction-red dotted).

Further, the graphical representation of the error MSA 10 days prediction for NWC-CHF path (low-mid-latitude) in Figure 5.46. Further, the statistical performance calculates to show the capability of the model predictor. The minimum value of the prediction error is -11.10 dB, the maximum value is 14.87 dB, the mean value is 2.88 dB, the standard deviation is 3.42 dB the error prediction shows in RMSE of 4.46 dB (NWC-CHF).

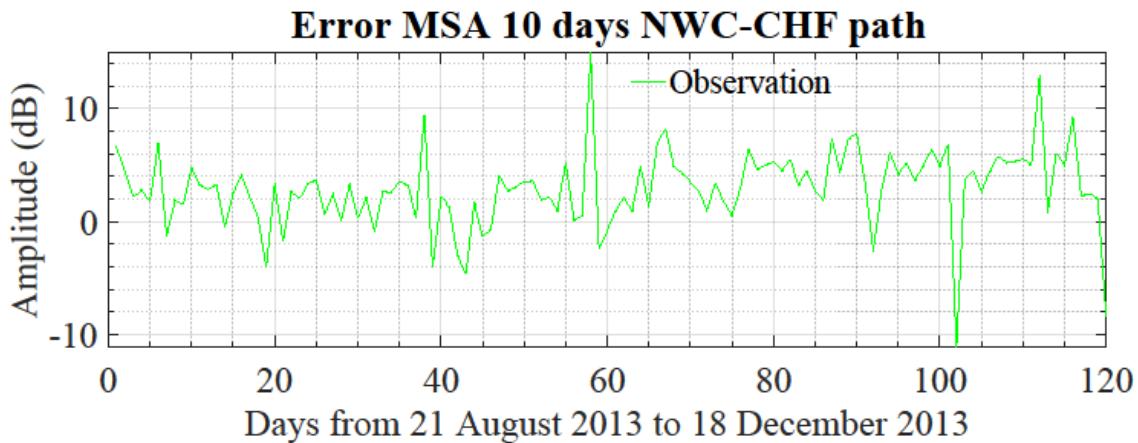


Figure 5.46: Error Multi Step (10 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves low-mid-latitude path over the time interval from 21 August 2013 to 18 December 2013.

Moreover, to examine the capability of our NARXNN DirRec model predictor, we use different datasets from different receiving station over the time interval between 15 March 2014 and 28 September 2015. The results are depicted in Figure 5.47. The blue curve denotes the actual VLF amplitude values and the red curve shows the predicted value. The constructed model successful for representing dynamic changing of VLF amplitude even for 10 days ahead as represented by Pearson correlation coefficient (r) of 0.0.683.

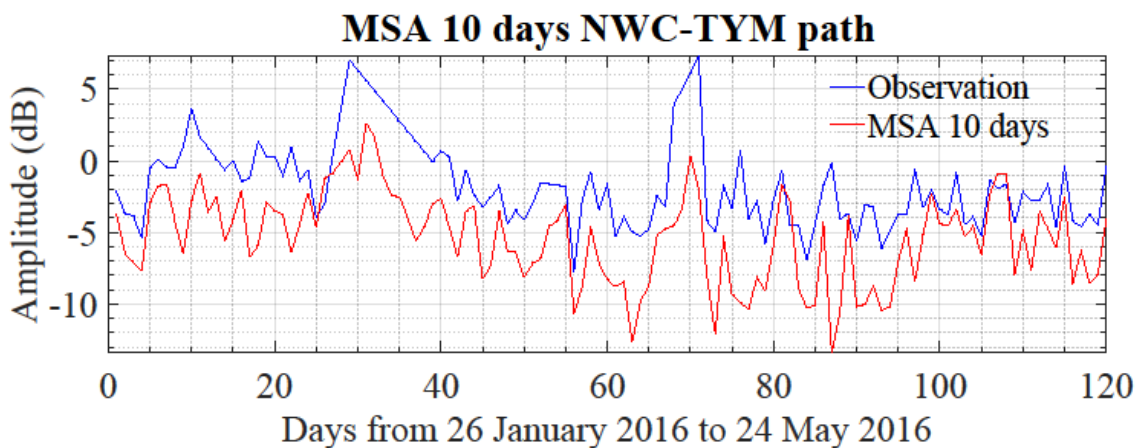


Figure 5.47: Multi Step (10 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves low-mid-latitude path with three days of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 26 January 2016 to 24 May 2016 (VLF observation-blue solid; Prediction-red dotted).

Further, the graphical representation of the error MSA 10 days prediction for NWC-TYM path (low-mid-latitude) in Figure 5.48. Further, the statistical performance calculates to

show the capability of the model predictor. The minimum value of the prediction error is -1.70 dB, the maximum value is 13.16 dB, the mean value is 3.77 dB, the standard deviation is 2.47 dB the error prediction shows in RMSE of 4.50 dB (NWC-TYM).

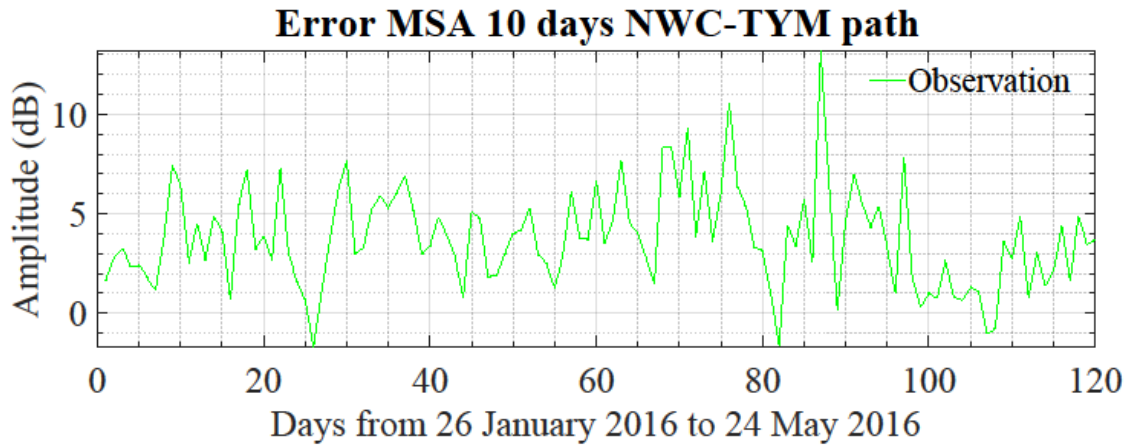


Figure 5.48: Error Multi Step (10 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves low-mid-latitude path over the time interval from 26 January 2016 to 24 May 2016.

5.3.2.2 Mid-Latitude VLF propagation path

To show the capability of NARXNN MSA prediction models in mid-latitude VLF propagation path, the curve plot of the observed data (blue curve) versus the predicted value of 5 days ahead in the 120 days data sets are shown in Figure 5.49 for NPM-CHF path. As a result, almost for all VLF paths of the predicted value suitably follow the observed value as indicated by a correlation coefficient is 0.84.

Further, the graphical representation of the error MSA 5 days prediction for NPM-CHF path (mid-latitude) in Figure 5.50. Further, the statistical performance calculates to show the capability of the model predictor. The minimum value of the prediction error is -12.99 dB, the maximum value is 8.62 dB, the mean value is 0.18 dB, the standard deviation is 3.13 dB the error prediction shows in RMSE of 3.12 dB (NPM-CHF).

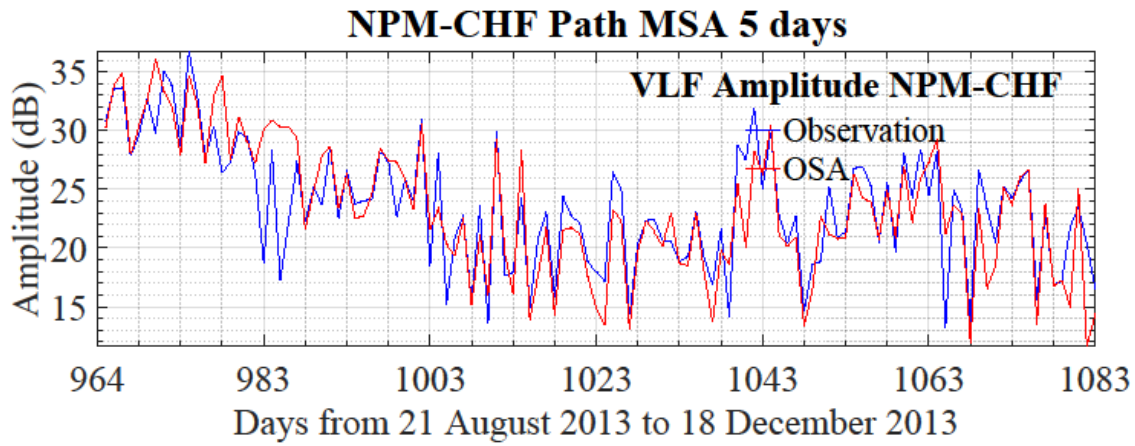


Figure 5.49: Multi Step (5 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves mid-latitude path with three days of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 21 August 2013 to 18 December 2013 (VLF observation-blue solid; Prediction-red dotted).

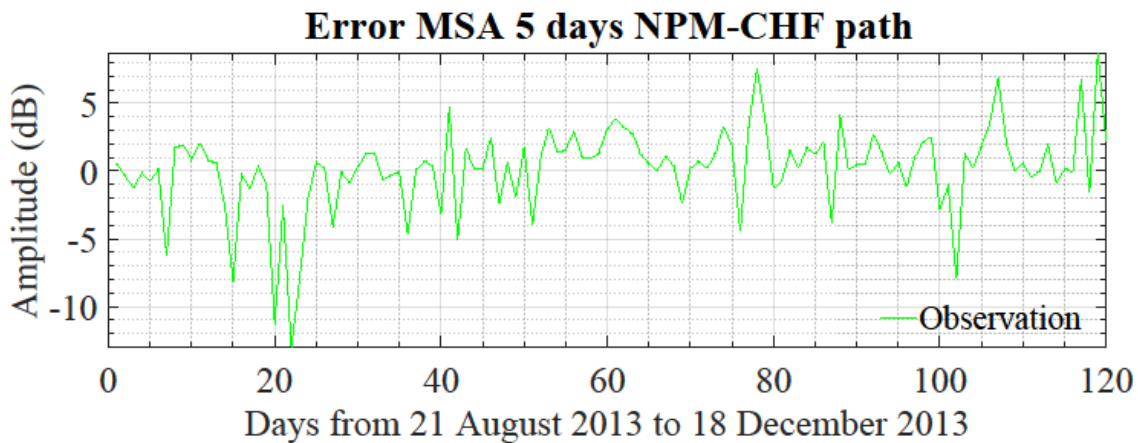


Figure 5.50: Error Multi Step (5 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves mid-latitude path over the time interval from 21 August 2013 to 18 December 2013 (VLF observation-blue solid; Prediction-red dotted).

Moreover, to examine the capability of our NARXNN DirRec model predictor, we use different datasets from different receiving station over the time interval between 15 March 2014 and 28 September 2015. The curve plot of the observed data (blue curve) versus the predicted value of 5 days ahead in the 120 days data sets are shown in Figure 5.51 for NPM-TYM path. As a result, almost for all VLF paths of the predicted value suitably follow the observed value as indicated by a correlation coefficient is 0.85.

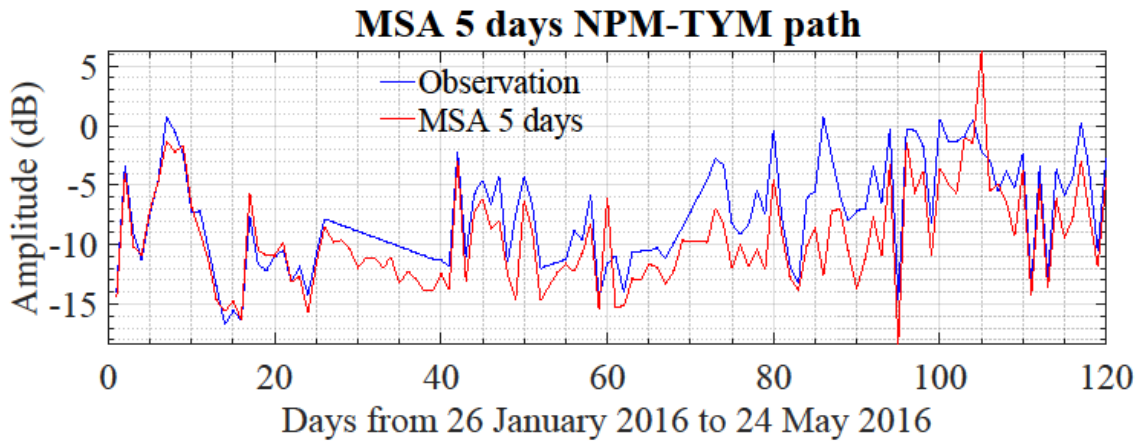


Figure 5.51: Multi Step (5 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves mid-latitude path with three days of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 26 January 2016 to 24 May 2016 (VLF observation-blue solid; Prediction-red dotted).

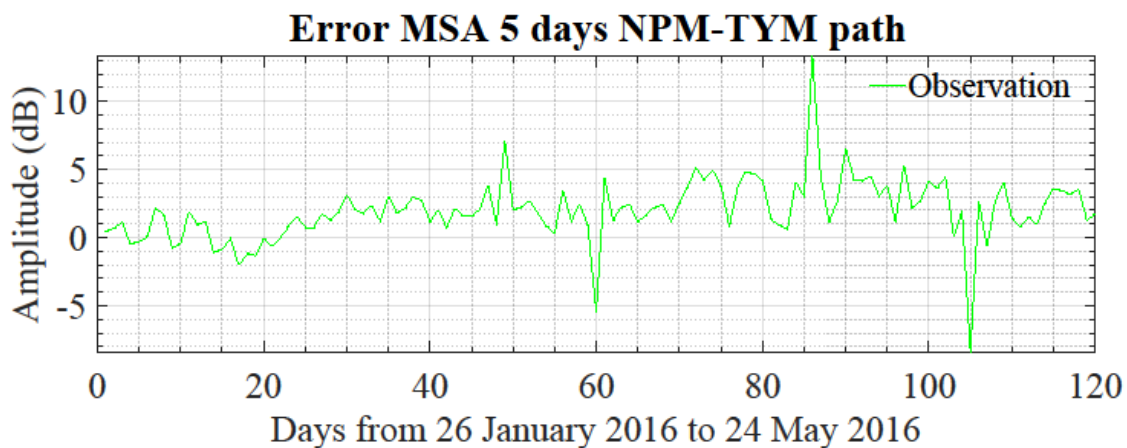


Figure 5.52: Error Multi Step (5 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves mid-latitude path over the time interval from 26 January 2016 to 24 May 2016.

Further, the graphical representation of the error MSA 5 days prediction for NPM-TYM path (mid-latitude) in Figure 5.52. Further, the statistical performance calculates to show the capability of the model predictor. The minimum value of the prediction error is -8.41 dB, the maximum value is 13.28 dB, the mean value is 1.96 dB, the standard deviation is 2.28 dB the error prediction shows in RMSE of 3.00 dB (NPM-CHF).

The same data sets are used to perform the 10 days ahead prediction of the daily nighttime of VLF electric field amplitude. Figure 5.53 shows the 10 days ahead prediction for mid-latitude (NPM-CHF). The blue curve denotes the actual VLF amplitude values and the red

curve shows the predicted value. The prediction performances for 10 days ahead are represented by a correlation coefficient of 0.73.

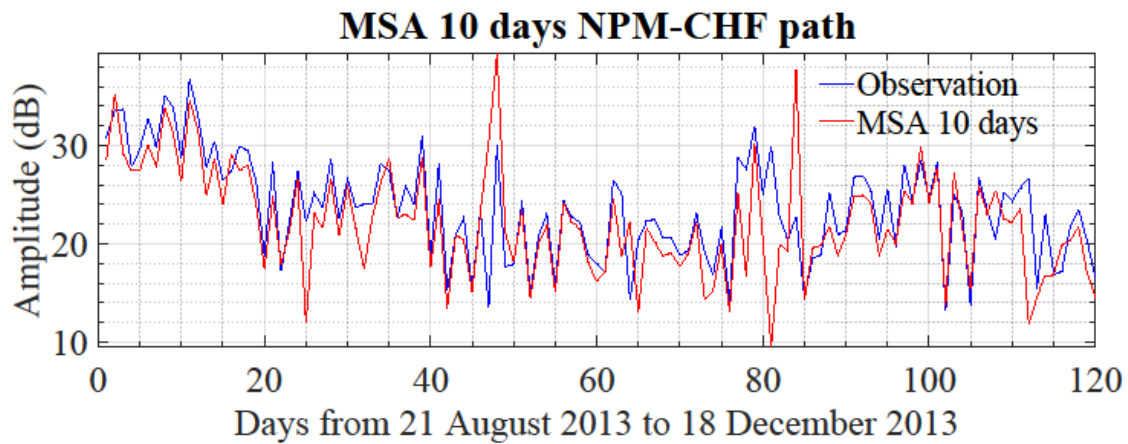


Figure 5.53: Multi Step (10 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves mid-latitude path with three days of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 21 August 2013 to 18 December 2013 (VLF observation-blue solid; Prediction-red dotted).

Further, the graphical representation of the error MSA 10 days prediction for NPM-CHF path (mid-latitude) in Figure 5.54. Further, the statistical performance calculates to show the capability of the model predictor. The minimum value of the prediction error is -16.69 dB, the maximum value is 20.28 dB, the mean value is 1.38 dB, the standard deviation is 3.97 dB the error prediction shows in RMSE of 4.20 dB (NPM-CHF).

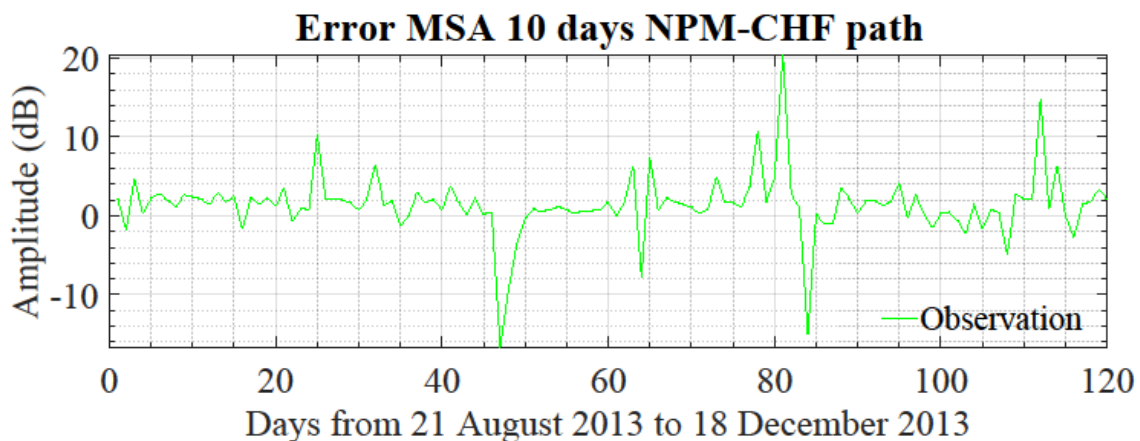


Figure 5.54: Error Multi Step (10 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves mid-latitude path over the time interval from 21 August 2013 to 18 December 2013

Moreover, to examine the capability of our NARXNN DirRec model predictor, we use different datasets from different receiving station over the time interval between 15 March 2014 and 28 September 2015. The results are depicted in Figure 5.55. The blue curve denotes the actual VLF amplitude values and the red curve shows the predicted value. The constructed model successful for representing dynamic changing of VLF amplitude even for 10 days ahead as represented by Pearson correlation coefficient (r) of 0.728.

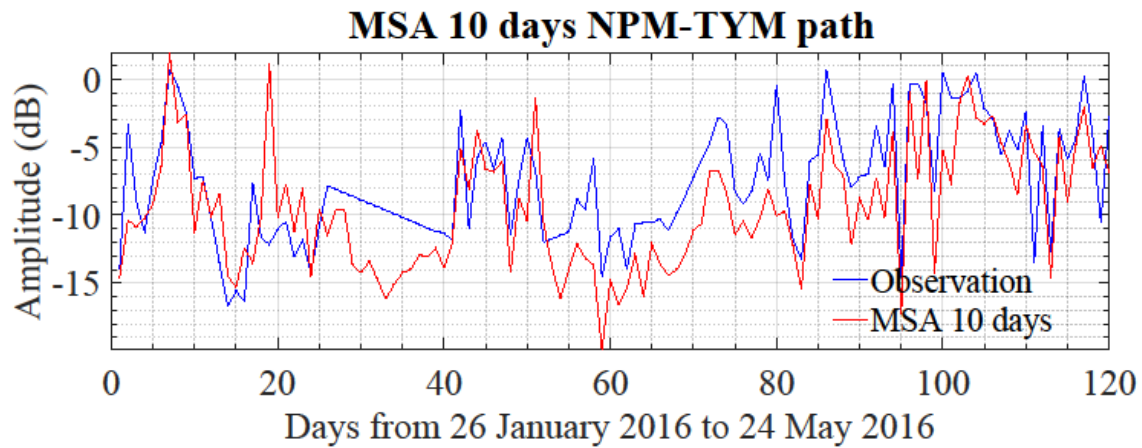


Figure 5.55: Multi Step (10 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves mid-latitude path with three days of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 26 January 2016 to 24 May 2016 (VLF observation-blue solid; Prediction-red dotted).

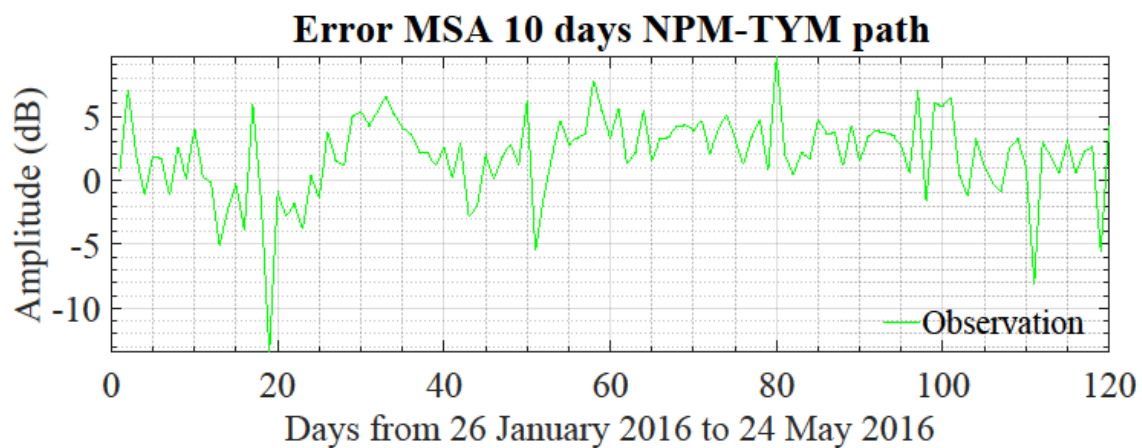


Figure 5.56: Error Multi Step (10 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves mid-latitude path over the time interval from 26 January 2016 to 24 May.

Further, the graphical representation of the error MSA 10 days prediction for NPM-TYM path (mid-latitude) in Figure 5.56. Further, the statistical performance calculates to show the capability of the model predictor. The minimum value of the prediction error is -13.39 dB, the maximum value is 9.59 dB, the mean value is 1.94 dB, the standard deviation is 3.23 dB the error prediction shows in RMSE of 3.76 dB (NPM-CHF).

5.3.2.3 High-Latitude VLF propagation path

The capability of NARXNN DirRec MSA prediction model in high-latitude VLF propagation path need to be examined, the blue curve represents the observed data and the red curve is defined as predicted value of 5 days ahead in the 120 days data sets. The result is shown in Figure 5.57 for NLK-CHF path. As a result, almost for all VLF paths of the predicted value suitably follow the observed value as indicated by a correlation coefficient is 0.80.

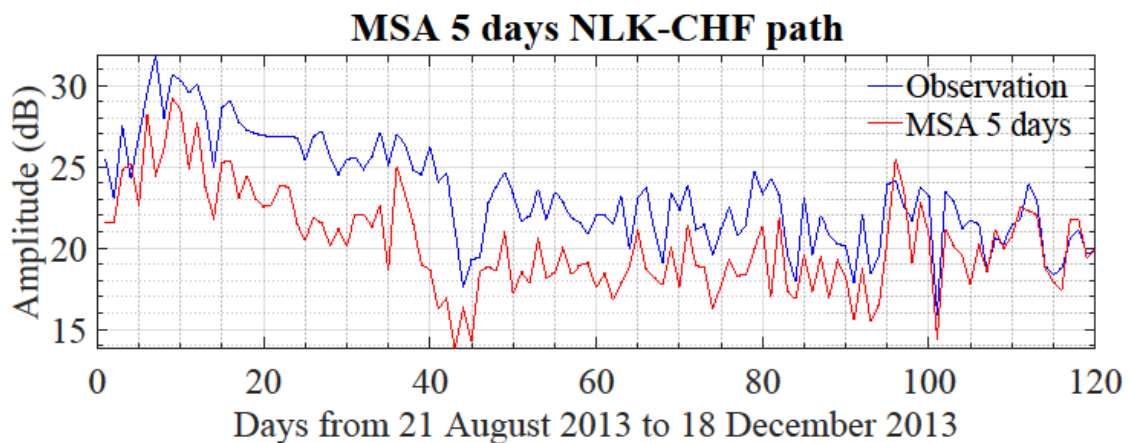


Figure 5.57: Multi Step (5 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves high-latitude path with three days of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 21 August 2013 to 18 December 2013 (VLF observation-blue solid; Prediction-red dotted).

Further, the graphical representation of the error MSA 5 days prediction for NLK-CHF path (high-latitude) in Figure 5.58. Further, the statistical performance calculates to show the capability of the model predictor. The minimum value of the prediction error is -1.31 dB, the maximum value is 7.69 dB, the mean value is 2.98 dB, the standard deviation is 1.97 dB the error prediction shows in RMSE of 3.57 dB (NLK-CHF).

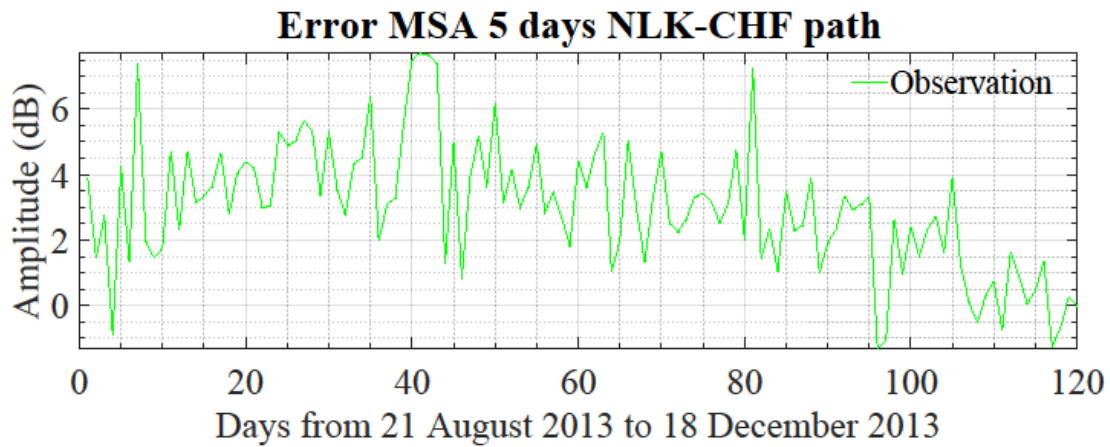


Figure 5.58: Error Multi Step (5 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves high-latitude path over the time interval from 21 August 2013 to 18 December 2013.

Moreover, to examine the capability of our NARXNN DirRec model predictor, we use different datasets from different receiving station over the time interval between 15 March 2014 and 28 September 2015. The blue curve represents the observed data and the red curve is defined as predicted value of 5 days ahead in the 120 days data sets are shown in Figure 5.59 for NLK-TYM path. As a result, almost for all VLF paths of the predicted value suitably follow the observed value as indicated by a correlation coefficient is 0.80.

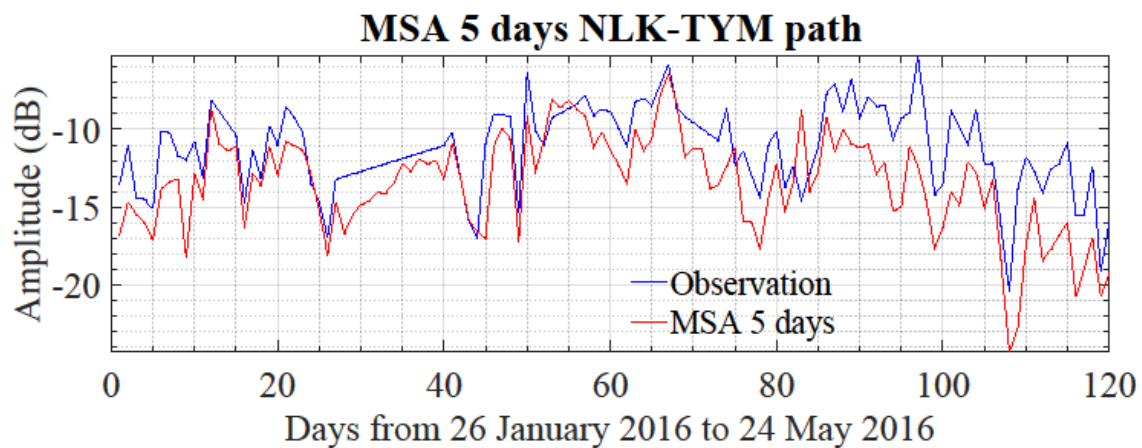


Figure 5.59: Multi Step (5 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves high-latitude path with three days of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 26 January 2016 to 24 May 2016 (VLF observation-blue solid; Prediction-red dotted).

Further, the graphical representation of the error MSA 5 days prediction for NLK-TYM path (high-latitude) in Figure 5.60. Further, the statistical performance calculates to show the capability of the model predictor. The minimum value of the prediction error is -5.85 dB, the maximum value is 8.90 dB, the mean value is 2.27 dB, the standard deviation is 1.93 dB the error prediction shows in RMSE of 2.97 dB (NLK-TYM).

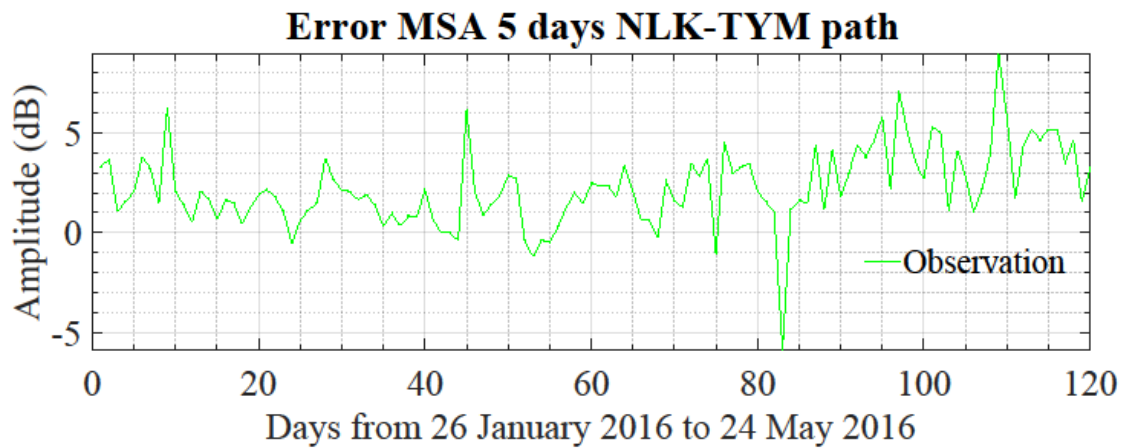


Figure 5.60: Error Multi Step (5 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves high-latitude path over the time interval from 26 January 2016 to 24 May 2016.

The same data sets are used to perform the 10 days ahead prediction of the daily nighttime of VLF electric field amplitude. Figure 5.61 shows the 10 days ahead prediction for mid-latitude (NLK-CHF). The blue curve denotes the actual VLF amplitude values and the red curve shows the predicted value. The prediction performances for 10 days ahead are represented by a correlation coefficient of 0.70.

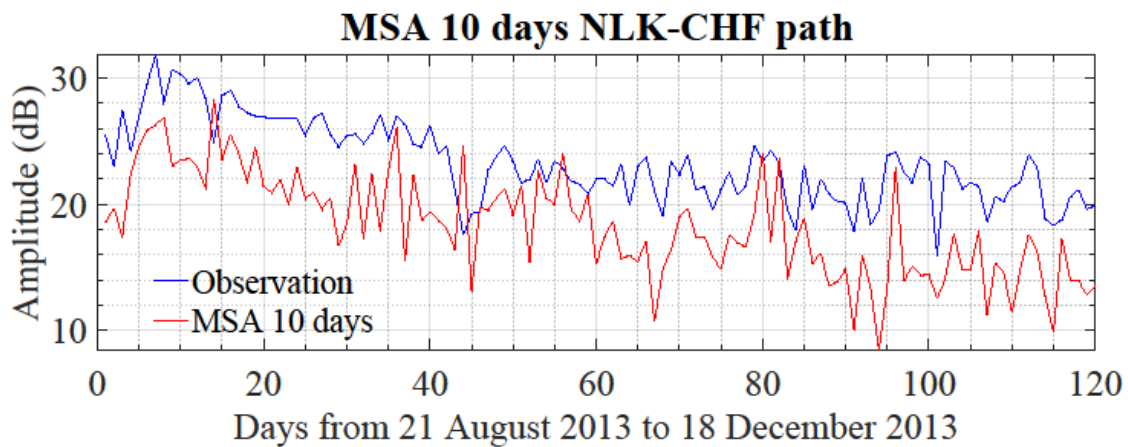


Figure 5.61: Multi Step (10 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves high-latitude path with

three days of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 21 August 2013 to 18 December 2013 (VLF observation-blue solid; Prediction-red dotted).

Further, the graphical representation of the error MSA 10 days prediction for NLK-CHF path (high-latitude) in Figure 5.62. Further, the statistical performance calculates to show the capability of the model predictor. The minimum value of the prediction error is -7.05 dB, the maximum value is 11.13 dB, the mean value is 5.02 dB, the standard deviation is 2.92 dB the error prediction shows in RMSE of 5.80 dB (NLK-CHF).

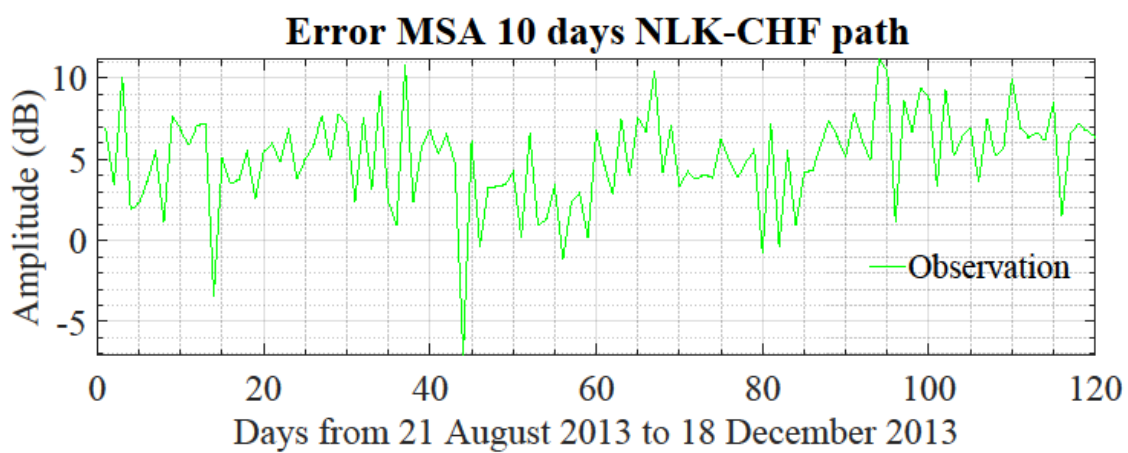


Figure 5.62: Error Multi Step (10 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves high-latitude path over the time interval from 21 August 2013 to 18 December 2013.

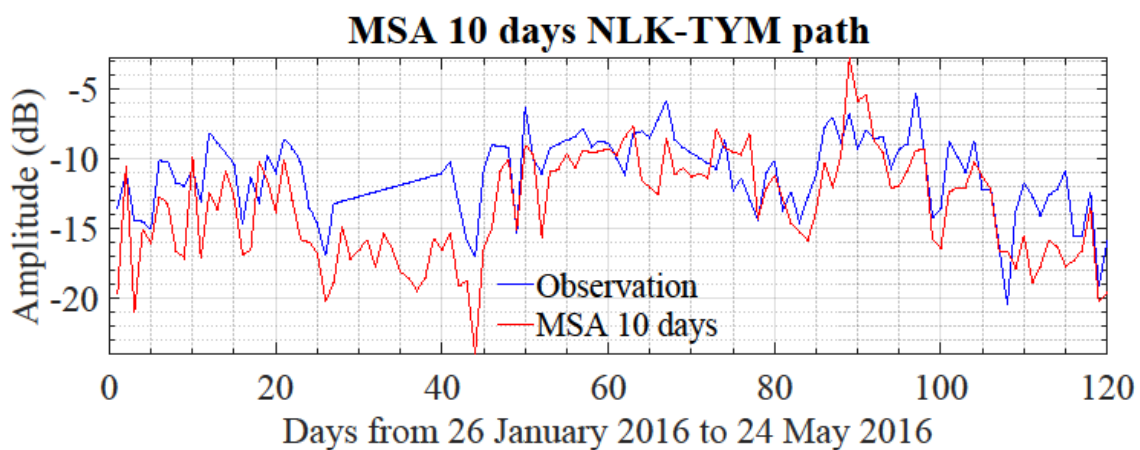


Figure 5.63: Multi Step (10 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves high-latitude path with three days of input-memory and two hundred neurons in the hidden layer by using LMANN algorithm over the time interval from 26 January 2016 to 24 May 2016 (VLF observation-blue solid; Prediction-red dotted).

Moreover, to examine the capability of our NARXNN DirRec model predictor, we use different datasets from different receiving station over the time interval between 15 March 2014 and 28 September 2015. The results are depicted in Figure 5.63. The blue curve denotes the actual VLF amplitude values and the red curve shows the predicted value. The constructed model successful for representing dynamic changing of VLF amplitude even for 10 days ahead as represented by Pearson correlation coefficient (r) of 0.725.

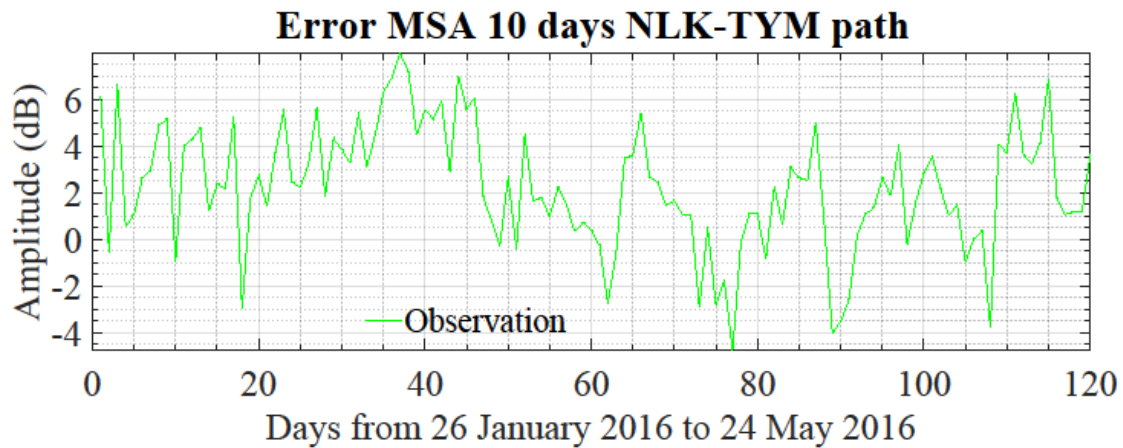


Figure 5.64: Error Multi Step (10 days) Ahead (MSA) predictions outside training period of NARXNN model of daily nighttime mean amplitude of VLF waves high-latitude path over the time interval from 26 January 2016 to 24 May.

Table 5.3: Fit Pearson's correlation coefficient (r) and root mean squared error (RMSE) of multi-step ahead prediction of daily VLF amplitude both Chofu and Tsuyama stations.

Path	Chofu				Tsuyama			
	MSA 5 days		MSA 10 days		MSA 5 days		MSA 10 days	
	r	RMSE	r	RMSE	r	RMSE	r	RMSE
Low-mid-latitude	0.790	2.60 dB	0.630	4.46 dB	0.801	3.57 dB	0.683	4.50 dB
Mid-latitude	0.840	3.12 dB	0.730	4.20 dB	0.850	3.00 dB	0.728	3.76 dB
High-latitude	0.800	3.57 dB	0.700	5.80 dB	0.800	2.97 dB	0.725	3.46 dB

Further, the graphical representation of the error MSA 10 days prediction for NLK-TYM path (high-latitude) in Figure 5.64. Further, the statistical performance calculates to show

the capability of the model predictor. The minimum value of the prediction error is -4.73 dB, the maximum value is 7.95 dB, the mean value is 2.27 dB, the standard deviation is 2.62 dB the error prediction shows in RMSE of 3.46 dB (NLK-TYM).

Table 5.3 summarize the performance of Chofu and Tsuyama for multi-step ahead VLF prediction model inside and outside the training periods. NARXNN MSA prediction model using LMANN algorithm with the input memory of 3 days before the given day and 200 neurons in the hidden layer is found to have a relatively good performance for different latitude paths. As shown in Table 5.3, Pearson's correlation coefficient of 5-day and 10-day ahead predictions both receiver station for different latitude greater than 0.630 and RMSE less than 5.80 dB.

5.4 Conclusions

In this chapter has presented for the first time a technique allowing the prediction of daily mean nighttime wave amplitude for subionospheric VLF propagation and identify relative significant physical parameters influencing in the nighttime D-region by using nonlinear autoregressive with exogenous input neural networks. The constructed model can predict both one step ahead and multi-step ahead prediction of VLF electric field amplitude on the high-, mid-, and low-mid-latitude paths based on the long-term experimental data.

Using constructed model from Chapter 5 for predicting different datasets from three VLF latitude paths and two different receiving station provide an important information to study the lower ionospheric conditions and the physical understanding about the parameters influencing in the built model. Moreover, the one-day VLF electric field amplitude has a good agreement with the observation (real) value with the average Pearson correlation coefficient both for all latitude paths and VLF station of 0.93 and root mean squared error of 1.33 dB over the training period. Further, the prediction performance is similarly good even outside the training period with r of 0.92 and RMSE of 1.60 dB.

Therefore, the average lower ionospheric properties along the high-, mid-, and low-mid-latitude paths are well represented by NARXNN nonlinear dynamic model. The constructed model will provide useful information for physical insight on the differing response of the nighttime lower ionosphere due to various different external forcings in the different latitude regions.

In the low-mid-latitude path the most controlling external forcing parameters to ionospheric variabilities are F10.7 index one and two days before the given day for Tsuyama

and Chofu stations respectively, Dst index one and two days before the given day for Tsuyama and Chofu stations respectively, and Kp index from previous day for both receiving stations. Further, in mid-latitude path atmospheric parameters have significant contribution than space weather effect as show by stratospheric temperature from the several days before and mesospheric temperature from previous day for both Chofu and Tsuyama. Solar activity as represent by F10.7 one day before the given day. Space weather effects is dominant external forcing parameter in high-latitude path. Dst index one day before the given day indicates as the most significant parameter contributed to the prediction model for both receiving station and followed by cosmic rays one and three days before the given day for Tsuyama and Chofu stations respectively. Total column ozone one and two days before the given day is also recognized as the significant parameter for Chofu and Tsuyama stations respectively.

Multi-step-ahead forecasting is still an open challenge in time series forecasting. Since the longer prediction horizon means the higher uncertainty, it is a very difficult task to predict the multi-steps for highly nonlinear VLF electric field amplitude. In Section 6.3 we attribute the performances of hybrid NARXNN and DirRec strategy for predicting five and ten days ahead of the daily nighttime mean value of VLF electric field amplitude for three different latitude paths.

The findings in this thesis support two previous studies which claim the superiority of NARXNN nonlinear modeling used for predicting VLF electric field amplitude in the mid-latitude path and DirRec strategy was implemented for multi-step ahead prediction of photovoltaic power [De Giorgi *et al.*, 2016]. The results of our study are relatively good with the Pearson correlation coefficient varies between 0.63 and 0.85. In addition, the root mean squared error (RMSE) for multi-step ahead prediction varies from 2.6 dB to 5.8 dB.

Further, the evaluation of the r and RMSE showed that all the datasets if the prediction horizon increases than the accuracy of the NARXNN DirRec strategy is decrease shown by decreasing r value and RMSE also increasing.

CHAPTER 6

6. Application for foF2 Data

The upper Earth's ionosphere is interesting issues to study because this region reflected radio waves to long distance and modifies the propagation of satellite signal. This region has been studied using vertical incidence sounding by sweep frequency ionosonde since 1932 [Appleton *et al.*, 1937]. Further, ionosonde from the bottom side and top side become a basic tool for studying the ionospheric conditions. The critical frequency (foF2) is the ordinary wave critical frequency of the highest stratification in the F-region (F2 layer). The success of high frequency (HF) communications depends largely on the ability to be able to predict the F2 region peak electron density. Moreover, foF2 is one of the most important parameters as far as communication coverage at a specific transmission frequency is concerned.

There have been several attempts in predicting the critical frequency (foF2), such as Nakamura *et al.* [2009] was developed a model to forecast ionospheric variations and storms at Kokubunji station by using a neural network. An empirical model to predict storm-time ionospheric foF2 from Haiku and Chongqing stations was developed by using support vector machine method [Ban *et al.*, 2011]. In addition, a general algorithm based on the neural network also use for predicting the ionospheric critical frequency of the F2 layer (foF2) in several ionosonde stations in China [Wang *et al.*, 2013].

Furthermore, in this chapter, we developed a nonlinear modeling for predicting one step ahead both hourly and daily ionospheric critical frequency in the F2 region (foF2) in Kokubunji ionosonde station with the geographic latitude 35° N and longitude 139° E by using NARXNN with the Levenberg-Marquardt algorithm. In this regard, we adapted several indices from Nakamura *et al.* [2009] for describing foF2 value such as, the day of the year (DOY)1, DOY2, sunspot number (SSN), F10.7, Dst, AE, and Kp indices. To find the significantly relative important parameters, we analyzed past foF2 in a period of 52 years from 1 January 1964 to 31 December 2016. Further, for the training propose we use 36 years datasets over the time interval between 1 January 1964 to 31 December 2000 and for prediction 16 years datasets with a range from 1 January 2001 to 31 December 2016.

6.1 Inputs/output Data Processing

The selection of the input datasets was influenced by the physical studies carried out to understand the physical meaning of the relationship between foF2 measured value changing and F2 layer condition. The data inputs datasets into the NARXNN were hourly and daily values of foF2, DOY1, DOY2, F10.7, Dst, AE, and Kp indices, SSN. Before the datasets used in the training and prediction, there was preprocessing data as explain in Section 5.2.

Table 6.1 shows the interval time of input dataset for foF2 prediction model both hourly and daily values. The input dataset for hourly foF2 values with the time interval between 1 January 1964 and 31 December 2016 (464,616 data points). The training period used a data from 1 January 1964 to 31 December 2000 (324,360 data points) and the prediction period between 1 January 2001 to 31 December 2016 (140,256 data points). Daily foF2 value has the same time interval from 1 January 1964 and 31 December 2016 (19,359 data points). Further, divided into two periods, between 1 January 1964 to 31 December 2000 (13,515 data points) is called training period and prediction period from 1 January 2001 to 31 December 2016 (5,844 data points).

Table 6.1: The interval data of foF2 prediction model for training and prediction periods

Description	Training period	Prediction period
Hourly foF2 prediction	01-01-1964 to 31-12-2000 (324,360 data points)	01-01-2001 to 31-12-2016 (140,256 data points)
Daily foF2 prediction	01-01-1964 to 31-12-2000 (13,515 data points)	01-01-2001 to 31-12-2016 (5,844 data points)

Figure 6.1 illustrates hourly values of the target and inputs of NARXNN for foF2 at Kokubunji station over the time interval from 1 January 1964 to 31 December 2016. Figure 6.1(a) shows the hourly foF2 at Kokubunji ionosonde station over fifty-three years. This variation appears to have some solar magnetic activity cycle which is nearly periodic 11-year change in the Sun's activity. It shows solar minimum and maximum effect appear in the critical frequency (foF2) measurement. Figure 6.1(b) and Figure 6.1(c) represent a combination of seasonal variation effect in one year. The hourly mean of solar radio flux at 10.7 cm (2800 MHz) data in solar flux units is shown Figure 6.1(d). Furthermore, Figure 6.1(e) shows hourly sunspot number measurement is provided by solar influences data analysis center (SIDC), Royal Observatory of Belgium, Brussels Belgium, Finally, Figure 6.1(f), Figure 6.1(g), and Figure 6.1(h) show Dst, AE, and Kp indices respectively. As seen from Figure 6.1, the temporal

dependences of the foF2, F10.7 index and sunspot number appear to have similar solar cycle dependences.

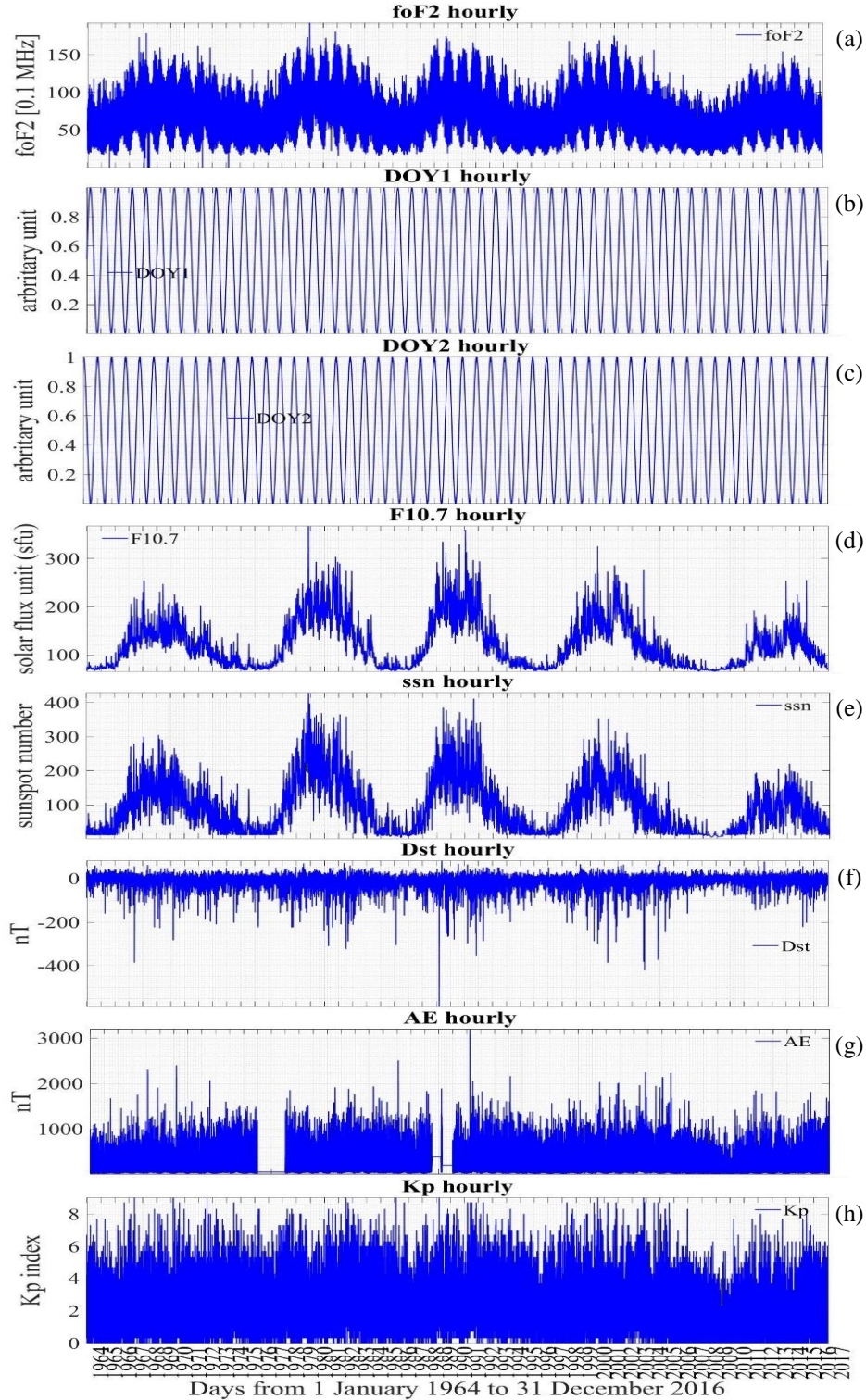


Figure 6.1: Hourly values of the target and inputs of NARXNN over the time interval from 1 January 1964 to 31 December 2016. (a) foF2 for Kokubunji ionosonde station as an input and target of the system (0.1 MHz), and six parameters as inputs to the system, (b) day of the year 1 (DOY1), (c) day of the year 2 (DOY2), (d) F10.7 index (sfu), (e) sunspot number (ssn), (f) Dst index (nT), (g) AE index (nT), and (h) Kp index.

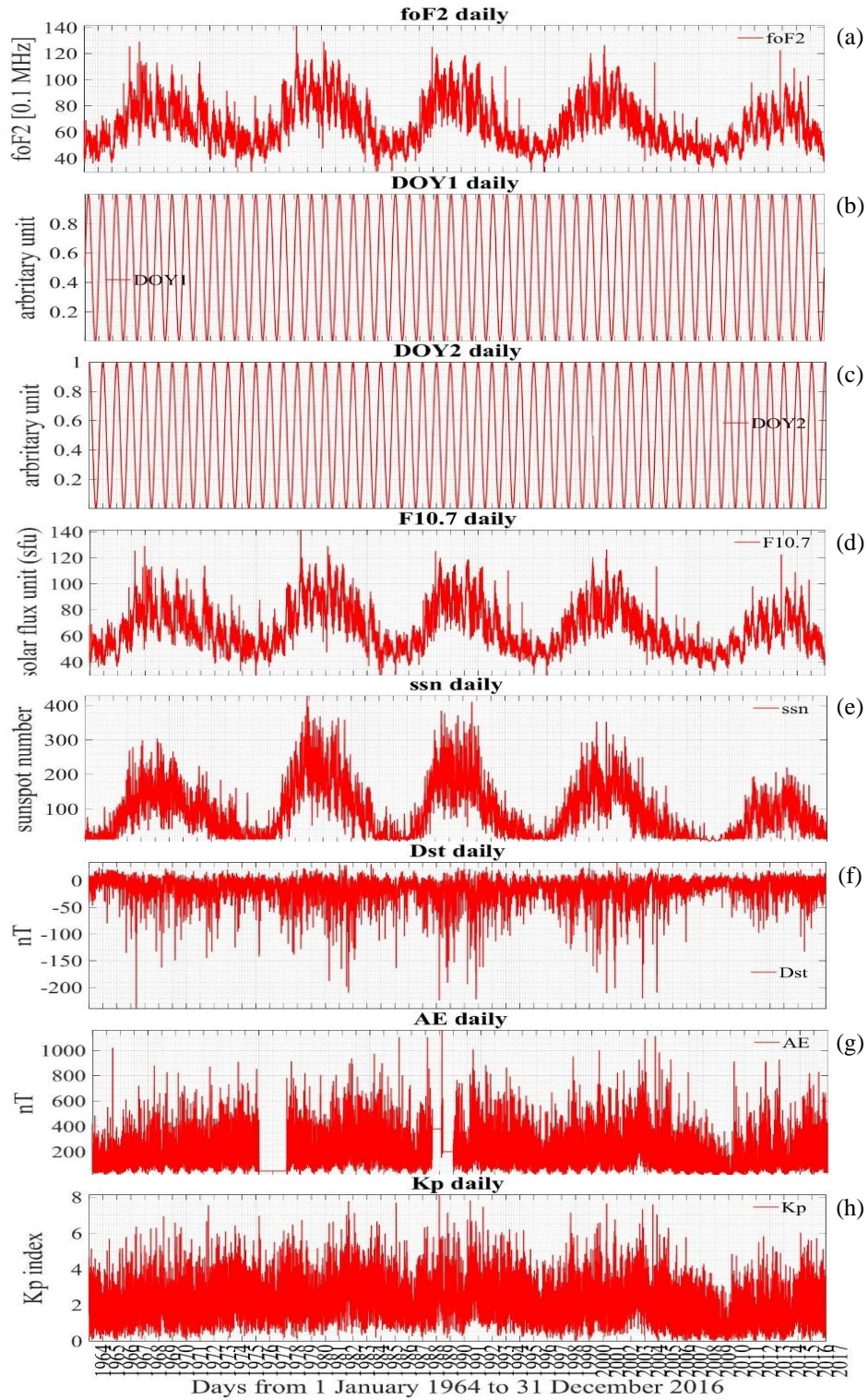


Figure 6.2: Daily values of the target and inputs of NARXNN over the time interval from 1 January 1964 to 31 December 2016. (a) foF2 for Kokubunji ionosonde station as an input and target of the system (0.1 MHz), and six parameters as inputs to the system, (b) day of the year 1 (DOY1), (c) day of the year 2 (DOY2), (d) F10.7 index (sfu), (e) sunspot number (ssn), (f) Dst index (nT), (g) AE index (nT), and (h) Kp index.

Figure 6.1 illustrates hourly foF2 with the time interval between 1 January 1964 and 31 December 2016. In addition to the graphical representation of the foF2, we use the statistical information to measure central tendency of the datasets. The minimum value is 0, the maximum value is 192, the mean value is 66.21, the standard deviation is 26.88.

Figure 6.2 shows daily values of the target and inputs of NARXNN for foF2 at Kokubunji station over the time interval from 1 January 1964 to 31 December 2016. Figure 6.2(a) - Figure 6.2(h) similar description with the hourly datasets. The foF2, F10.7 and sunspot number as seen in Figure 6.2 more clearly indicate that temporal dependence between those parameters is similar to solar cycle effect.

6.2 Architecture of foF2 NARXNN Modeling

The constructed network consists of three main layers namely input layer, hidden layer, and output layer. First, the input layer is representing of the current and previous stage of inputs and outputs. And then, the input layer value will be fed into the hidden layer. Second, the hidden layer is representing of one or few number of neurons resulting in a nonlinear mapping of the connected combination of the weights values from the input layer. Third, the output layer is representing of the connected combination of the weights values from the hidden layer. The dynamical order of inputs and outputs and number of neurons in each layer are pre-determined. *Haykin* [1998] is represent a few methods for determining of those values. A appropriate training algorithm and performance measure has to be selected. Finally, the category of the nonlinear map also requires to be determined.

The input and target values need for pre- and post-processing in order to obtain a valid training set. Further, those processes consist of normalization of the input and target data into values with the interval from -1 to 1 , normalization values for the inputs and targets to obtain zero mean and unity variance and removal of constant inputs and outputs. For evaluating the performance of a network, performance criteria are chosen to observe the error between desired responses (original data) and calculated outputs (prediction).

In this thesis, the NARXNN, where the exogenous inputs (u) are DOY1, DOY2, F10.7, Dst, AE, and Kp indices, SSN have been implemented using Levenberg Marquardt Neural Network (LMANN) Algorithm. The foF2 was used as the output (y). Both inputs and output have a memory of two hours/days from the given hour/day. The NARXNN consist ten neurones in the hidden layer each utilizing tangent sigmoid activation function. The NARXNN

model for one-step ahead prediction has been diagrammatically shown in Figure 6.3. Further, the NARXNN model for the temporal dependence of the foF2 can be given by equation (6.1).

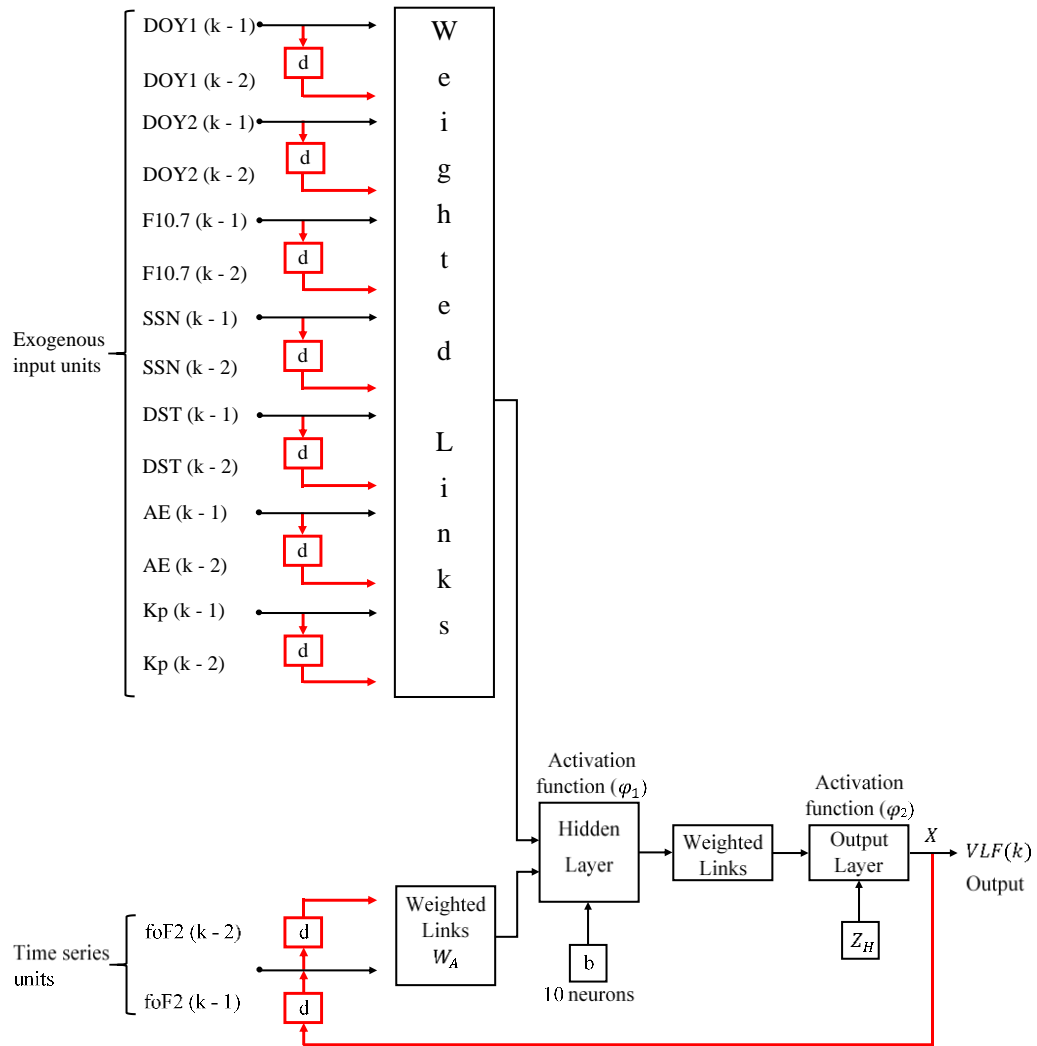


Figure 6.3: NARXNN architecture of the critical frequency in the F2 region (foF2) prediction.

$$\begin{aligned}
foF2(k+1) = F[& (foF2(k), \dots, foF2(k - d_{foF2} \\
& + 1); DOY1(k), \dots, DOY1(k - d_{DOY1} \\
& + 1); DOY2(k), \dots, DOY2(k - d_{DOY2} \\
& + 1); F10.7(k), \dots, F10.7(k - d_{F10.7} \\
& + 1); SSN(k), \dots, SSN(k - d_{SSN} \\
& + 1); Dst(k), \dots, Dst(k - d_{Dst} \\
& + 1); AE(k), \dots, AE(k - d_{AE} \\
& + 1); Kp(k), \dots, Kp(k - d_{Kp} + 1)]
\end{aligned} \tag{6.1}$$

where $F[.]$ denotes a nonlinear tangent sigmoid function and d_{foF2} , d_{DOY1} , d_{DOY2} , $d_{F10.7}$, d_{SSN} , d_{Dst} , d_{AE} , and d_{Kp} are maximum lagged time in step foF2 for the output, and DOY1, DOY2, F10.7, Dst, AE, and Kp indices, SSN for inputs, respectively. Moreover, nonlinear tangent sigmoid function given by equation (6.2)

$$F(x) = \frac{1}{1 + \exp(-x)} - \frac{1}{2} \tag{6.2}$$

The NARXNN performances can be evaluated using root mean square error (RMSE). Further, the RMSE can range between 0 and ∞ and it should be noted that the minimum RMSE value is the best network performance in prediction. In order to calculate the RSME, the difference between estimated and actual values are each squared and then averaged over the sample period and then rooted. RMSE can be given by equation (6.3)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (error_{sim}(i))^2} \tag{6.3}$$

where n is the number of data points (one point per day), and $error_{sim}(i)$ is the error difference (in dB) between the foF2 from observation and the model prediction at the certain data point (day i).

The Pearson's correlation coefficient (CC) is also used to provide information about the NARXNN performances. Further, the CC can take a range of values from +1 to -1. A value

of 0 indicates that there is no association between the two variables. A value greater than 0 indicates a positive association; that is, as the value of one variable increases, so does the value of the other variable. A value less than 0 indicates a negative association; that is, as the value of one variable increases, the value of the other variable decreases. The CC of prediction data over the observed values is calculated using (6.4)

$$CC = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}} \quad (6.4)$$

where

- N the total number of predicted value
- x_i the observed data
- y_i the predicted value
- \bar{x} the mean value of x_i
- \bar{y} the mean value of y_i

6.3 Prediction Results

In this section, the results of the prediction computed with two different datasets are presented along with the statistical performance. The NARXNN model was estimated for the datasets both hourly and daily critical frequency (foF2) with the time period from 1 January 1964 to 31 December 2016 from ionosonde measurement station in Kokubunji, Tokyo, Japan. Further, the relative importance of the parameter will represent physical sources which have a contribution to the prediction model. Moreover, the prediction results inside the training period over the time interval from 1 January 1964 to 31 December 2000 and outside the training period with the time interval between 1 January 2011 to 31 December 2016 are illustrated graphically in this section. In addition, statistical performance of the error prediction will be shown to judge the performance of the NARXNN constructed model.

6.3.1 Hourly foF2 Prediction

The NARXNN model equation shows the relative importance parameters which are influenced by the constructed model as represented with the coefficients. The highest coefficient means the contributions to the prediction of foF2 also significant and vice versa

[Garson, 1991; Goh, 1995]. The complete NARXNN model equation for predicting foF2 hourly is represented in equation (6.5). Further, the first four significant terms for critical frequency hourly are described here. The first influential factor is the foF2 with one hour before the given hour $\{foF2(k-1)\}$ as represented by the coefficient of 11.603. The F10.7 with one hour before the given hour $\{F10.7(k-1)\}$ becomes the second significant factor with the coefficient of 5.012. The third influential factor is the F10.7 with two hours $\{F10.7(k-2)\}$ before the given hour with a weighting coefficient of 4.661. The last one is Kp index with one hour before the given hour $\{Kp(k-1)\}$ as indicated by the coefficient of 4.580. Further, the relative importance term will graphically illustrate in Figure 6.4. The four of the most relative importance parameter will normalized in percentage within those parameters.

$$foF2(k) = \begin{bmatrix} -11.603 foF2(k-1) + 1.234 foF2(k-2) - 0.239 doy1(k-1) \\ + 0.200 doy1(k-2) - 3.994 doy2(k-1) - 0.727 doy2(k-2) \\ + 5.012 f10.7(k-1) + 4.661 f10.7(k-2) - 3.244 ssn(k-1) \\ + 0.286 ssn(k-2) + 0.167 Dst(k-1) + 1.703 Dst(k-2) \\ - 0.152 AE(k-1) + 2.839 AE(k-2) - 4.580 kp(k-1) \\ + 3.393 kp(k-2) - 0.642 \end{bmatrix} \quad (6.5)$$

The foF2 variable is the most significant contribution to the prediction model of hourly value. It indicates variabilities of the foF2 value influence foF2 prediction, which is highly correlated with the consecutive hour. This high correlation indicates that the foF2 value does not vary a great deal from the value of the hour before if no strong external forcing exists. The F10.7 index is the most significant external forcing with large coefficients that contributed in the prediction model. The critical frequency in F2-region is the maximum frequency which can be reflected by the F2 layer at vertical incidence and is proportional to the square root of the maximum electron density in F2-region. Solar EUV radiation heating the upper ionosphere region during the daytime enhanced electron production rate. The high electron density in the F2-region followed by higher critical frequency (foF2) [Ikubanni and Adeniyi, 2012]. Maximum electron density increase with increasing solar activity and in F2 region associated to the altitude variation of the effective loss rate for electrons and to diffusion effects of maximum electron production [Reid, 1972]. The foF2 in the ionosphere saturates due to high electron density at high solar activity. The significant variations of foF2 with solar activity due to several factors such as neutral winds, thermospheric wind, dynamo electric field and the varying Sun-Earth distance due to the Earth's elliptic orbit round the Sun. The vertical magnetic field ($E \times$ magnetic field (B) force which is induced by atmospheric tidal forcing, is due to the conjunction of an eastward-westward electric field (E). Further, foF2 has a good correlation in

average with F10.7 index for moderate solar activity and a strong solar activity dependence of foF2 around equinoxes than solstices shown by *Ikubanni et al.* [2013]. *Brum et al.*, [2011] mentioned foF2 has strong dependence with solar activity, increase with increasing solar activity and varying with seasons. In moderate solar activity, foF2 is strongest dependence with solar activity because the decreasing phase of the solar cycle compared to high and low solar activity. Moreover, *Yang and Chen*, [2009] have been found the neutral winds are stronger in solstices than in equinoxes and the increasing strength of the winds tends to cause depletion of electron density with consequently a drop in foF2 measurement during solstices. The neutral wind is affected by the temperature and the pressure of the ionosphere. This is closely related with the location of projection of the sunlight on Earth. While in equinox the sunlight is projected on the vicinity of the equator, the sunlight is straight projected on North or South Hemispheres in solstices. Thus, in solstices the temperature is higher and the neutral winds are more active correspondingly. Consequently, the couplings between North and South Hemispheres are much stronger in solstice seasons. On the other hand, when North Hemisphere is in winter, Earth is on the location that is closest to Sun. By comparison, Earth is farthest from Sun when North Hemisphere is in summer. Therefore, the ionosphere is much more active in winter than in summer, and the coupling in winter is stronger than in summer. One hour up to two-hour time lag represent to a small trend of F10.7 index is highly correlated with local fluctuation of foF2 measurement.

The ionospheric response associated to geomagnetic storms as seen in Kp index has been found by *Klimenko et al.*, [2017]. Energy from ring current region is propagated to F-region along magnetic fields line in the form of heat of the electron gas, electron being thermalized by Coulomb collisions with ring current particles. The Kp index showed high values due to geomagnetic storms occurred on 26-29 September 2011 followed by foF2 disturbances at Irkutsk and Kaliningrad. The effects were caused by an increase in the $n(O)/n(N_2)$ ratio. *Lobzin and Pavlov* [2002] were found the NmF2 negative disturbance occurrence probability dependence shows the strong positive Kp tendency with the increase in the maximum absolute value of the NmF2 negative disturbance amplitude. One-hour time lag for Kp index shows the effect of geomagnetic activity instantaneously to the ionospheric disturbances.

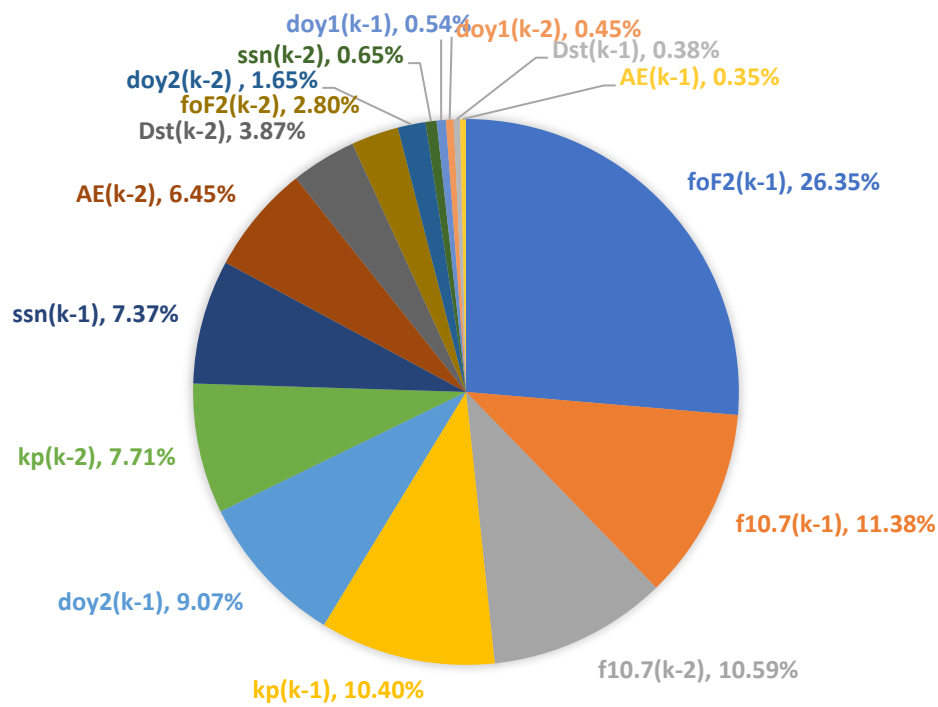


Figure 6.4: The relative significant parameter in hourly prediction of critical frequency (foF2) in Kokubunji station.

The NARXNN model with two hours of input-memory and ten neurons in the hidden layer using LMANN algorithm is used to predict the hourly of foF2. The fitted model result is represented by Figure 6.5. The fitted model (inside the training period) with the time interval from 1 January 1964 to 31 December 2000 show in red curve and observation in blue curve. The constructed model worked well for prediction as represented by Pearson correlation coefficient (r) is 0.9605.

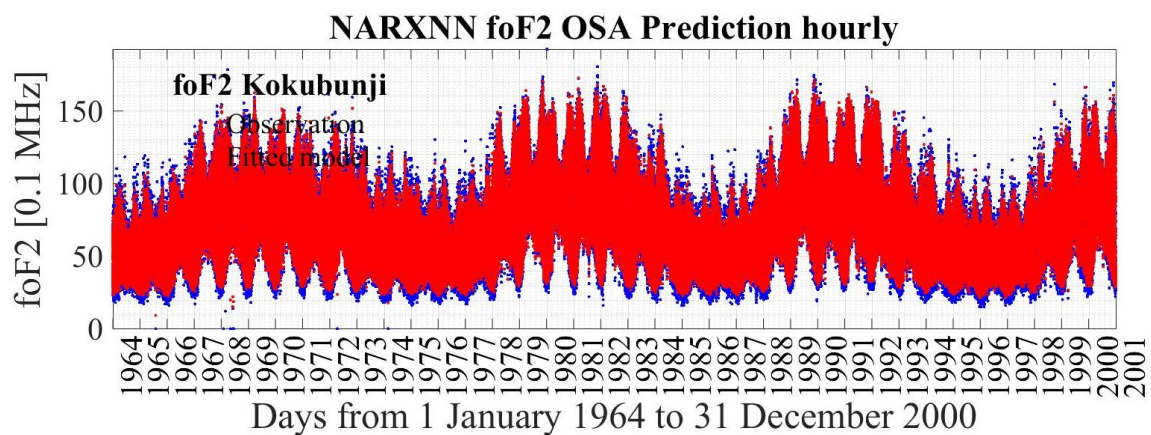


Figure 6.5: The fitted model predictions of NARX NN model of hourly foF2 Kokubunji station with two-day of input-memory and ten neurons in the hidden layer by using LMANN algorithm over the time interval from 1 January 1964 to 31 December 2000.

Furthermore, the prediction error of fitted model inside the training period for 324,360 hours data point is shown graphically in Figure 6.6. Further, the statistical performance calculates to show the capability of the model predictor. The minimum value of the prediction error is -98.41 [0.1 MHz], the maximum value is 86.11 [0.1 MHz], the mean value is 0.04 [0.1 MHz], the standard deviation is 7.73 [0.1 MHz], the error prediction shows in RMSE of 7.72 [0.1 MHz].

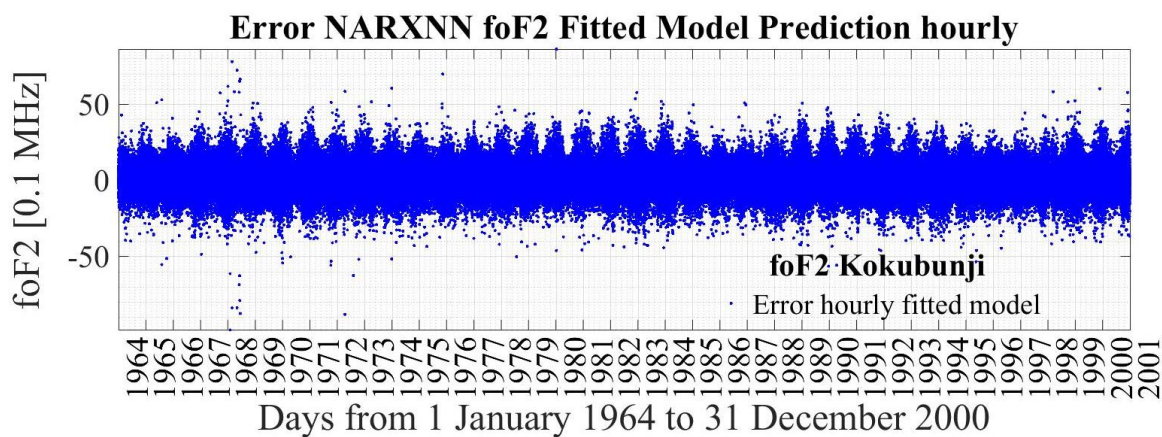


Figure 6.6: Error fitted model predictions of NARX NN model of hourly foF2 Kokubunji station with two-day of input-memory and ten neurons in the hidden layer by using LMANN algorithm over the time interval from 1 January 1964 to 31 December 2000.

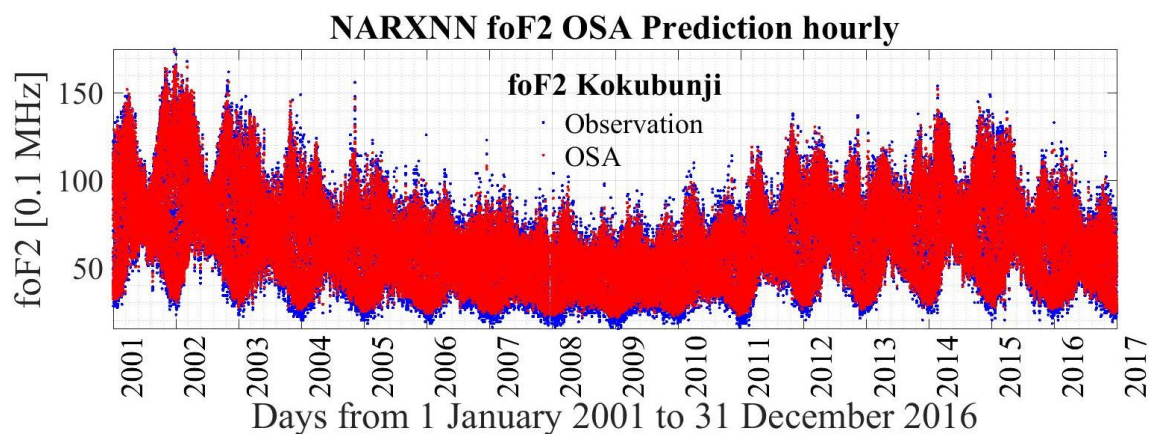


Figure 6.7: One Step (1 day) Ahead (OSA) predictions of NARX NN model of hourly foF2 Kokubunji station with two-day of input-memory and ten neurons in the hidden layer by using LMANN algorithm over the time interval from 1 January 2001 to 31 December 2016.

Figure 6.7 shows the OSA prediction with the data set outside the training period. The correlation coefficient for the prediction remains high value as represented by r of 0.9548.

Furthermore, the prediction error of fitted model inside the training period for 140,256 hours data point is shown graphically in Figure 6.8. Further, the statistical performance calculates to show the capability of the model predictor. The minimum value of the prediction error is -56.60 [0.1 MHz], the maximum value is 67.31 [0.1 MHz], the mean value is -0.11 [0.1 MHz], the standard deviation is 7.20 [0.1 MHz], the error prediction shows in RMSE of 7.20 [0.1 MHz].

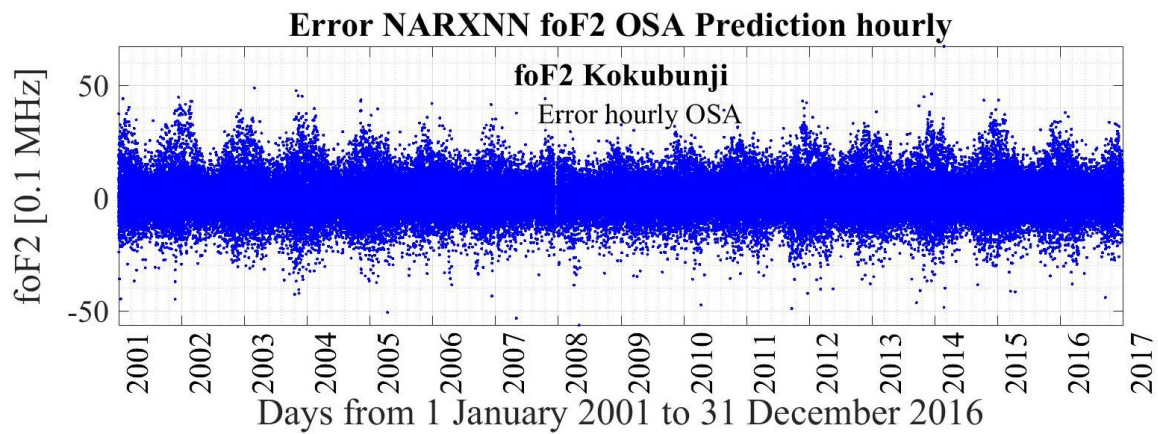


Figure 6.8: Error one Step (1 day) Ahead (OSA) predictions of NARX NN model of hourly foF2 Kokubunji station with two-day of input-memory and ten neurons in the hidden layer by using LMANN algorithm over the time interval from 1 January 2001 to 31 December 2016.

6.3.2 Daily foF2 Prediction

Moreover, equation (6.6) shows the complete NARXNN model equation for predicting foF2 daily. Further, the first four significant terms for critical frequency daily are described here. The first influential factor is the foF2 with one day before the given day $\{foF2(k-1)\}$ as represented by the coefficient of 7.271. The F10.7 with one day before the given day $\{F10.7(k-1)\}$ becomes the second significant factor with the coefficient of 6.335. The third influential factor is the AE index with one day $\{AE(k-1)\}$ before the given day with a weighting coefficient of 6.119. The last one is DOY1 with two days before the given day $\{DOY1(k-2)\}$ as indicated by the coefficient of 3.578. In addition, the relative importance term will graphically illustrate in Figure 6.4. The four of the most relative importance parameter will normalized in percentage within those parameters.

$$foF2(k) = \begin{bmatrix} 7.271 foF2(k-1) - 2.784 foF2(k-2) - 0.912 doy1(k-1) \\ -3.578 doy1(k-2) - 0.198 doy2(k-1) - 0.242 doy2(k-2) \\ -6.335 f10.7(k-1) - 0.097 f10.7(k-2) + 1.941 ssn(k-1) \\ -1.045 ssn(k-2) + 2.698 Dst(k-1) + 1.166 Dst(k-2) \\ +6.119 AE(k-1) - 1.241 AE(k-2) + 0.802 kp(k-1) \\ -1.444 kp(k-2) - 2.496 \end{bmatrix} \quad (6.6)$$

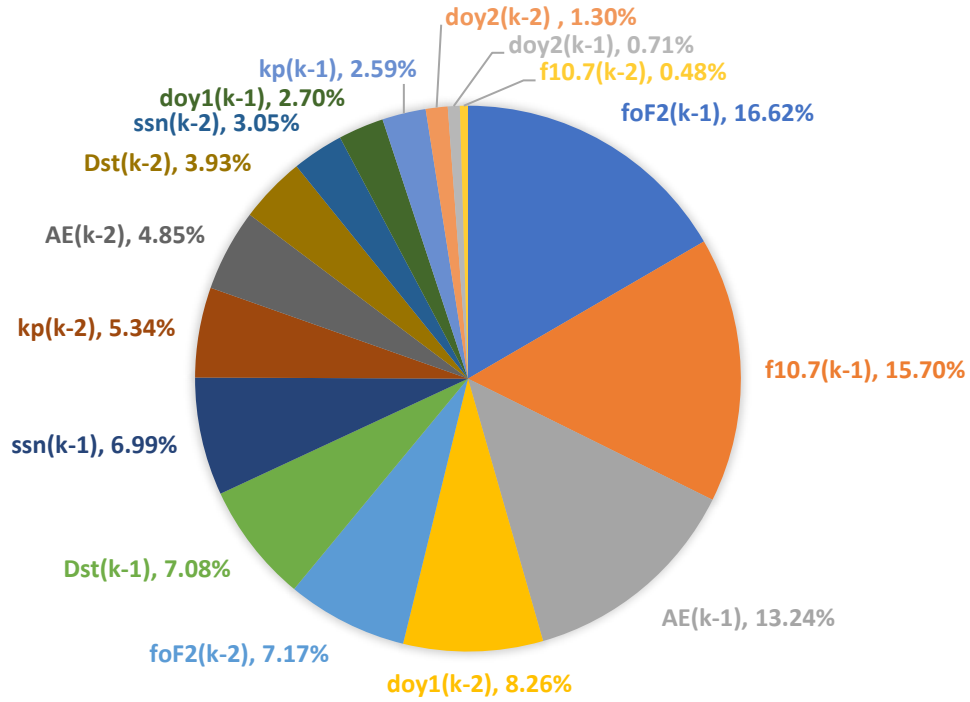


Figure 6.9: The relative significant parameter in daily prediction of critical frequency (foF2) in Kokubunji station.

The foF2 variable is the most significant contribution to the prediction model of daily value. It indicates variabilities of the foF2 value influence foF2 prediction, which is highly correlated with the consecutive day. This high correlation indicates that the foF2 value does not vary a great deal from the value of the day before if no strong external forcing exists. The F10.7 index is the most significant external forcing with large coefficients that contributed in the prediction model. As mentioned before that solar EUV radiation heating the upper ionosphere region during the daytime enhanced electron production rate. *Ikubanni and Adeniyi*, [2012] have been found that high electron density in the F2-region followed by higher critical frequency (foF2). In addition, *Reid*, [1972] shown maximum electron density increase with increasing solar activity and in F2 region associated to the altitude variation of the effective loss rate for electrons and to diffusion effects of maximum electron production. Further, foF2

has a good correlation in average with F10.7 index for moderate solar activity and a strong solar activity dependence of foF2 around equinoxes than solstices shown by *Ikubanni et al.* [2013]. Moreover, *Brum et al.*, [2011] mentioned foF2 has strong dependence with solar activity, increase with increasing solar activity and varying with seasons. In moderate solar activity, foF2 is strongest dependence with solar activity because the decreasing phase of the solar cycle compared to high and low solar activity. Moreover, *Yang and Chen*, [2009] have been found the neutral winds are stronger in solstices than in equinoxes and the increasing strength of the winds tends to cause depletion of electron density with consequently a drop in foF2 measurement during solstices. One day time lag represents to accumulation of small trend of F10.7 in the hourly value from a hour up to two hour time delay.

The geomagnetic activities are recognized as the second most significant external forcing to foF2 prediction. Ionospheric responses to intensive geomagnetic activity are highly complex and have a degree of variability. However, negative ionospheric disturbances associated to the negative storm effects are caused by a neutral composition disturbance zone (decrease of O/N₂) generated during geomagnetic storms at auroral latitudes and then transported to lower latitudes by the disturbed thermospheric wind circulation produced by Joule heating [*Buonsanto*, 1999; *Huang et al.*, 2005]. The composition of neutral changes the balance between electron production and loss rates and effected in decreasing NmF2. At middle-latitude ionosphere has been studied by *Huang et al.*, [2005] that the positive effects occurred a few hours after the storm sudden commencements which may be attributable to a manifestation of atmospheric disturbances (large-scale gravity waves) which are launched from auroral zone during storms and propagate to middle-latitude. Moreover, *Tsagouri et al.* [2000] has been investigated middle latitude ionospheric disturbances associated to geomagnetic storms using foF2 observations data and the results show the positive storm effects at mid-latitude have a correlation with foF2 measurement and changing in Dst and AE indices. The positive storm effects based on Prolss model are expected and attributed to traveling atmospheric disturbances (TADs) and highly correlated to AE. This may be expected since TADs are pulse-like atmospheric perturbations just triggered by a sudden energy release and then moving in the form of global and circumpolar front [*Prolss*, 1995]. Moreover, a good correlation shown by *Trísková* [1994] between foF2 value and AE index in few hours' time lag also with Dst index over 12 hour time lag. The energy from ring current region is propagated to F-region along magnetic fields line in the form of heat of the electron gas, electron being thermalized by Coulomb collisions with ring current particles. The ionospheric response associated to geomagnetic storms as seen in AE index has been found by the high values due

to geomagnetic storms occurred on 26-29 September 2011 followed by foF2 disturbances at Irkutsk and Kaliningrad. The effects were caused by an increase in the $n(O)/n(N_2)$ ratio.

The day of the year is the fourth significant parameter that contributed in the prediction model. DOY represent seasonal variation and *Ikubanni et al.*, [2013] have been investigated that foF2 variations dependence with seasonal and solar activity. Moreover, linear dependence has been identified between critical frequency and season [*Ikubanni and Adeniyi*, 2012]. The seasonal variation of foF2 may be due to the change in the neutral gas composition. The densities of these gases vary significantly across seasons due to the meridional winds irrespective of solar activity [*Stubbe*, 1975].

The NARXNN model with two days of input-memory and ten neurons in the hidden layer using LMANN algorithm is used to predict the daily of foF2. The fitted model result is represented by Figure 6.10. The fitted model (inside the training period) with the time interval from 1 January 1964 to 31 December 2000 show in red curve and observation in blue curve. The constructed model worked well for prediction as represented by Pearson correlation coefficient (r) is 0.9630.

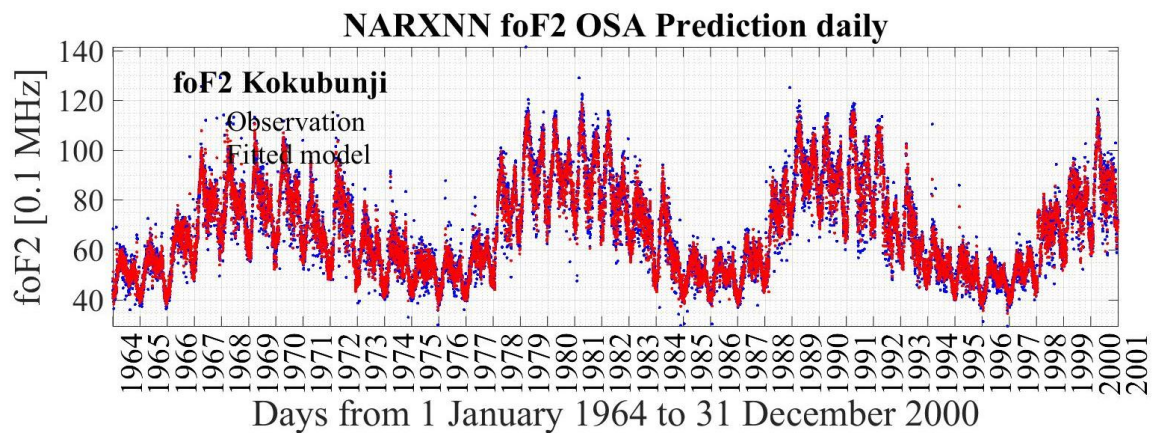


Figure 6.10: The fitted model predictions of NARX NN model of daily foF2 Kokubunji station with two-day of input-memory and ten neurons in the hidden layer by using LMANN algorithm over the time interval from 1 January 1964 to 31 December 2000.

Furthermore, the prediction error of fitted model inside the training period for 13,515 days data point is shown graphically in Figure 6.11. Further, the statistical performance calculates to show the capability of the model predictor. The minimum value of the prediction error is -32.32 [0.1 MHz], the maximum value is 56.90 [0.1 MHz], the mean value is 0.33 [0.1 MHz], the standard deviation is 4.91 [0.1 MHz], the error prediction shows in RMSE of 4.93 [0.1 MHz].

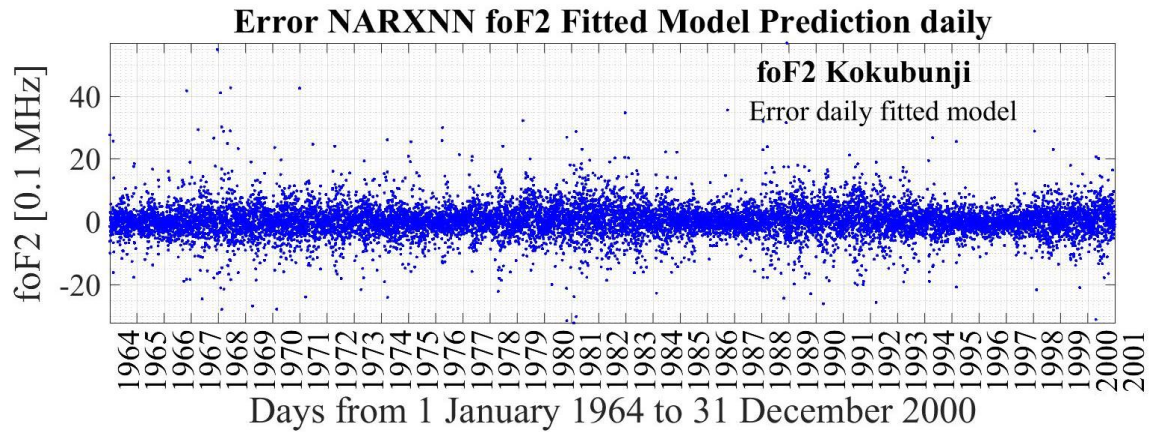


Figure 6.11: Error fitted model predictions of NARX NN model of daily foF2 Kokubunji station with two-day of input-memory and ten neurons in the hidden layer by using LMANN algorithm over the time interval from 1 January 1964 to 31 December 2000.

Figure 6.12 shows the OSA prediction with the data set outside the training period. The correlation coefficient for the prediction remains high value as represented by r of 0.9620.

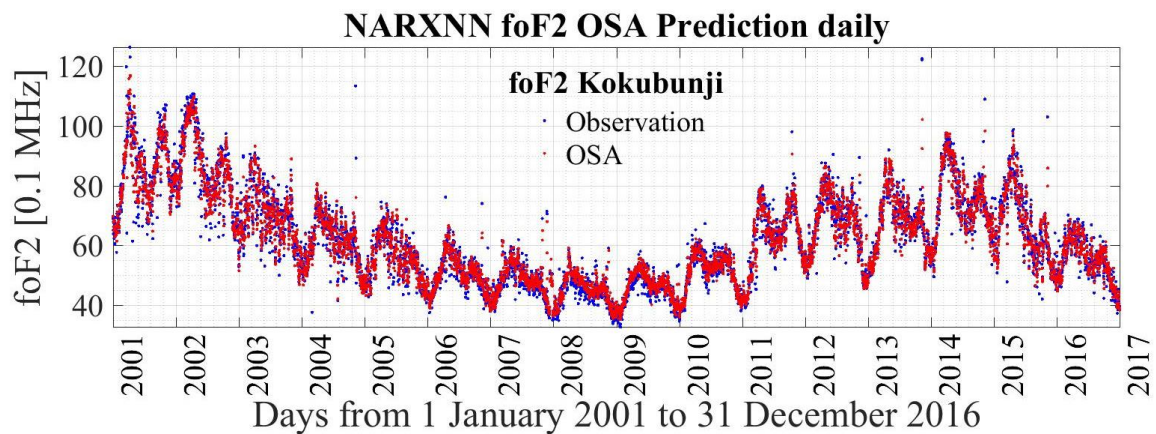


Figure 6.12: One Step (1 day) Ahead (OSA) predictions of NARX NN model of daily foF2 Kokubunji station with two-day of input-memory and ten neurons in the hidden layer by using LMANN algorithm over the time interval from 1 January 2001 to 31 December 2016.

Furthermore, the prediction error of fitted model inside the training period for 5,844 hours data point is shown graphically in Figure 6.13. Further, the statistical performance calculates to show the capability of the model predictor. The minimum value of the prediction error is -29.97 [0.1 MHz], the maximum value is 54.34 [0.1 MHz], the mean value is 0.04 [0.1 MHz], the standard deviation is 4.26 [0.1 MHz], the error prediction shows in RMSE of 4.26 [0.1 MHz].

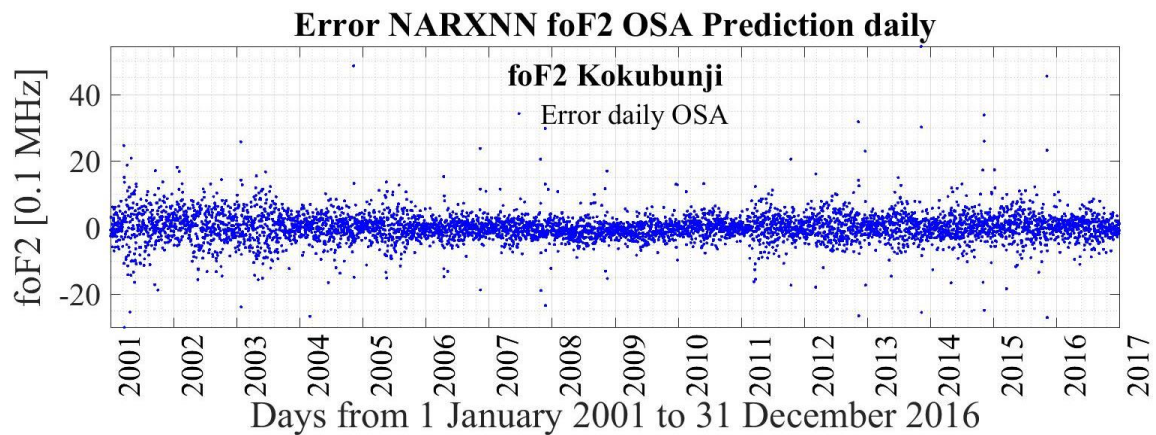


Figure 6.13: Error one Step (1 day) Ahead (OSA) predictions of NARX NN model of daily foF2 Kokubunji station with two-day of input-memory and ten neurons in the hidden layer by using LMANN algorithm over the time interval from 1 January 2001 to 31 December 2016.

Table 6.2 summarize the performance of hourly and daily foF2 prediction model inside and outside the training periods. NARXNN prediction model using LMANN algorithm with the input memory of 2 days before the given day and 10 neurons in the hidden layer is found to have a good performance both inside and outside training period. As shown in Table 6.2, Pearson's correlation coefficient both hourly and daily values greater than 0.95 and RMSE less than 0.78 MHz.

Table 6.2: Fit Pearson's correlation coefficient (r) and root mean squared error (RMSE) of foF2 hourly and daily predictions.

Value	Fitted model		OSA	
	r	RMSE	r	RMSE
Hourly	0.961	0.77 MHz	0.955	0.72 MHz
Daily	0.963	0.49 MHz	0.962	0.42 MHz

6.4 Conclusions

In this chapter, the NARXNN dynamic nonlinear model was extended to work with completely different datasets. Based on one of the most important parameters for communication purpose at specific transmission frequency is called foF2. Further, the success of high frequency (HF) communications depends largely on the ability to be able to predict the F2 region peak electron density. We used this parameter as an output prediction by using the proposed model both for one step ahead hourly and daily prediction.

Applying NARXNN dynamic nonlinear model for very long-term and different datasets lead to better training and prediction results shown in the model performance as represented by Pearson correlation coefficient and root mean squared error. Furthermore, OSA daily foF2 prediction has excellent performance with the r inside the training period of 0.9630 and RMSE of 0.49 MHz for 13,515 days of datasets. Moreover, for hourly foF2 prediction also has a high correlation coefficient of 0.9605 and RMSE of 0.77 MHz for 324,336 hours of datasets. In addition, outside the training period or real prediction, the proposed model also shows similarly good with r of 0.9620 and RMSE of 0.42 MHz for 5,844 days of datasets and r of 0.9548 and RMSE of 0.72 MHz for 140,256 of datasets. Further, the solar flux index from the previous day is recognized as the most significant contribution for external forcing to foF2 variation both hourly and daily values. The geomagnetic activities as represented by Kp and AE indices from previous day for hourly and daily value respectively are recognized as the significant external forcing to foF2 prediction.

The proposed NARXNN with the Levenberg-Marquardt algorithm is robust under totally changing the datasets. The average of upper ionospheric properties in Kokubunji, Tokyo, Japan is well represented by the constructed model. The relative significant parameters will provide useful information for physical insight on the dynamic response of the F2 region due to different external sources.

CHAPTER 7

7. Conclusions and Future Works

7.1 Conclusions

The aim of this thesis is to develop a nonlinear model system identification for predicting ionospheric conditions through the radio wave propagation and remote sensing using external sources both from below and above the ionosphere and to study ionospheric properties of a quantitative contribution from external forcings. Further, a novel implementation of a nonlinear autoregressive with exogenous input neural networks (NARXNN) has been presented. The approach is employed due to great capability in learning the dynamics of nonlinear systems and the capabilities to manage the system complexity. The performance of the proposed model is evaluated to predict the daily nighttime mean value of VLF electric field amplitude. In addition, in the F-region NARXNN model is used to predict the critical frequency in the F2 region (foF2) value for the long-term time interval.

Nonlinear system identification technique based on NARX structure with neural network approach has been considered. Three different training algorithms namely Bayesian regulation, Levenberg-Marquardt, and scale conjugate gradient algorithms to identify underlying highly complex relationships based on exogenous inputs-output have been discussing in Chapter 3. The influence number of time delay selection and neuron size in the hidden layer on the behaviour of three different training algorithm has been examined and the most suitable value for each variable chosen for NARXN dynamic modeling. Thus, based on the results among the three algorithms (BRANN, LMANN, and SCG) we consider that the LMANN algorithm with delay time three days before the given day and two hundred neurons in the hidden layer is found to have the best performance, with the largest correlation and the lowest prediction error. Moreover, the performance of the LMANN algorithm does not vary significantly with increasing number of neurons in the hidden layer and input memory, which enable us to maintain the high prediction performance with a relatively small number of neurons in the hidden layer and input memory that allows minimizing the computational cost is shown in Chapter 4.

Further, using the best NARXNN configuration for predicting different datasets from three latitude paths and two different receiving VLF station were tested to analyze the proposed model ability to predict in advance have been performed in Chapter 5. In addition, the NARXNN model equation provides relative important information to study the lower ionospheric conditions and the physical understanding of the parameters influencing in the built model. The constructed model clearly outperforms the persistence model even for the small horizon, which indicates that considering recent past observations is a better approach to perform prediction.

The constructed NARXNN model indicates in the low-mid-latitude path that VLF amplitude, F10.7, Dst, and Kp indices are the four of the most relative important contribution to VLF amplitude variation respectively. Further, stratospheric temperature, F10.7 index, VLF amplitude, and mesospheric temperature have been indicated the most four relative significant influencing VLF amplitude variation in mid-latitude path respectively. Finally, the most four relative important in the high-latitude path is Dst index, cosmic ray, AE index, and total column ozone respectively.

One day ahead VLF electric field amplitude prediction has a good agreement with the observation (real) value with the average Pearson correlation coefficient (r) both for all latitude paths and VLF station of 0.93 and root mean squared error of 1.33 dB over the training period. Further, the prediction performance is similarly good even outside the training period with r of 0.92 and RMSE of 1.60 dB. Therefore, the average lower ionospheric properties along the high-, mid-, and low-mid-latitude paths are well represented by NARXNN nonlinear dynamic model.

In the low-mid-latitude path, the most controlling external forcing parameters to ionospheric variabilities are F10.7 index one and two days before the given day for Tsuyama and Chofu stations respectively, Dst index one and two days before the given day for Tsuyama and Chofu stations respectively, and Kp index from the previous day for both receiving stations. Further, in mid-latitude path atmospheric parameters have significant contribution than space weather effect as shown by stratospheric temperature from the several days before and mesospheric temperature from the previous day for both Chofu and Tsuyama. Solar activity as represents by F10.7 one day before the given day. Space weather effects are dominant external forcing parameter in the high-latitude path. Dst index one day before the given day indicates as the most significant parameter contributed to the prediction model for both receiving stations and followed by cosmic rays one and three days before the given day for Tsuyama and Chofu stations respectively. Total column ozone one and two days before the given day are also recognized as the significant parameter for Chofu and Tsuyama stations respectively.

Multi-step-ahead forecasting is still an open challenge in time series forecasting. Since the longer prediction horizon means the higher uncertainty, it is a very difficult task to predict the multi-steps. Moreover, the prediction error increases nonlinearly with the increasing of step number ahead. The new technique of hybrid NARXNN and DirRec strategy is developed for the first time. The performances of hybrid NARXNN and DirRec strategy for predicting five and ten days ahead of the daily nighttime mean value of VLF electric field amplitude for three different latitude paths are evaluated. The results of our study are relatively good with the r varies between 0.63 and 0.85. In addition, the RMSE for multi-step ahead prediction varies from 2.6 dB to 5.8 dB. Further, the evaluation of the r and RMSE are showed the prediction horizon increases followed by decreases of the prediction accuracy.

The NARXNN dynamic nonlinear model was extended to perform with completely different datasets of hourly and daily foF2 from ionosonde measurement in Kokubunji station as evaluated in Chapter 6. The NARXNN dynamic nonlinear model for very long-term and different datasets lead to better training and prediction results shown in the model performance as represented with r and RMSE. Furthermore, OSA daily foF2 prediction has excellent performance with the r inside the training period of 0.9630 and RMSE of 0.49 MHz for 13,515 days of datasets. Moreover, for hourly foF2 prediction also has the high correlation coefficient of 0.9605 and RMSE of 0.77 MHz for 324,360 hours of datasets. In addition, outside the training period or real prediction, the proposed model also shows similarly good with r of 0.9620 and RMSE of 0.42 MHz for 5,844 days of datasets and r of 0.9548 and RMSE of 0.72 MHz for 140,256 of datasets. Thus, the proposed NARXNN with time delay two days before the given day and ten neurons in the hidden layer with the Levenberg-Marquardt algorithm is robust under totally changing the datasets. The average of F2 region properties around Kokubunji, Tokyo, Japan is well represented by the constructed model.

The solar flux index from the previous day is recognized as the most significant contribution for the external forcing to foF2 variation both hourly and daily values. The geomagnetic activities as represented by Kp and AE indices from the previous day for hourly and daily value respectively are recognized as the significant external forcing to foF2 prediction.

In the middle-latitude, lower and upper ionosphere coupling can be explained by the significant contribution of nonlinear modeling for VLF electric field amplitude (D-region) and critical frequency in F2-region (F2-region). The most common and significant external forcing both of two layers is solar radio flux at 10.7 cm (F10.7) index with one day before the given day. The result shows solar activity as represented by F10.7 index has the instant influence to the lower and upper ionosphere conditions.

7.2 Future Works

Any successful research effort invariably leaves further topics to be explored. A few items which are not covered in this thesis but which would be potentially exciting areas for future research are explained below.

First, there are many different paths and receiving stations of VLF transmitter-receiver wave propagation operated by UEC's VLF networks which allowing to study lower ionospheric conditions. However, in this thesis only the long path was modeled and evaluated using NARXNN. A short path such as JJI-CHF, JJY-CHF or domestic path which consist of ground wave or one hop propagation only and to study the other contribution of physical sources for understanding of coupling mechanisms to the lower ionosphere from the atmosphere and the Sun.

Second, the limitation of NARXNN is the significant contribution only represent by relative importance contribution. Other nonlinear system identification approaches can also be applied to predict VLF amplitude and foF2 which has the exact physical contribution represent with error reduction ratio (ERR) such as nonlinear autoregressive moving average with exogenous input (NARMAX). Comparison of this other approach may help to understand more about physical contribution for studying ionosphere conditions.

References

- Abdu, M. A., & Angreji, P. D. (1974). The role played by ozone in the lower D region electron density variations in winter. *Journal of Geophysical Research*, 79(4), 649–657. <https://doi.org/10.1029/JA079i004p00649>
- Abdu, M. A., Batista, I. S., Piazza, L. R., & Massambani, O. (1981). Magnetic storm associated enhanced particle precipitation in the South Atlantic Anomaly: Evidence from VLF phase measurements. *Journal of Geophysical Research: Space Physics*, 86(A9), 7533–7542. <https://doi.org/10.1029/JA086iA09p07533>
- Afraimovich, E. L., Palamartchouk, K. S., & Perevalova, N. P. (1998). GPS radio interferometry of travelling ionospheric disturbances. *Journal of Atmospheric and Solar-Terrestrial Physics*, 60(12), 1205–1223. [https://doi.org/10.1016/S1364-6826\(98\)00074-1](https://doi.org/10.1016/S1364-6826(98)00074-1)
- Akala, A. O., Doherty, P. H., Carrano, C. S., Valladares, C. E., & Groves, K. M. (2012). Impacts of ionospheric scintillations on GPS receivers intended for equatorial aviation applications. *Radio Science*, 47(4), 1–11. <https://doi.org/10.1029/2012RS004995>
- Alken, P., Maute, A., & Richmond, A. D. (2017). The F₂-Region Gravity and Pressure Gradient Current Systems: A Review. *Space Science Reviews*, 206(1–4), 451–469. <https://doi.org/10.1007/s11214-016-0266-z>
- Appleton, E. V., Naismith, R., & Ingram, L. J. (1937). British Radio Observations during the Second International Polar Year 1932-33. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 236(764), 191–259. <https://doi.org/10.1098/rsta.1937.0002>
- Arain, M., Ayala, H., & Ansari, M. (2012). Nonlinear System Identification Using Neural Network. *Emerging Trends and Applications in ...*, 122–131. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-28962-0_13
- Armstrong, W. C. (1983). Recent advances from studies of the Trimpf effect. *Antarct. J. of USA*, 18, 281–283.
- Åström, K. J., & Wittenmark, B. (2008). *Adaptive Control*. Dover Publications. Retrieved from https://books.google.co.jp/books?id=L0m_CR-IK24C
- Ayala, H. V. H., & Coelho, L. D. S. (2016). Cascaded evolutionary algorithm for nonlinear system identification based on correlation functions and radial basis functions neural networks. *Mechanical Systems and Signal Processing*, 68–69, 378–393. <https://doi.org/10.1016/j.ymssp.2015.05.022>

- Ban, P. P., Sun, S. J., Chen, C., & Zhao, Z. W. (2011). Forecasting of low-latitude storm-time ionospheric foF2 using support vector machine. *Radio Science*, 46(6), 1–9. <https://doi.org/10.1029/2010RS004633>
- Banks, P. M., & Kockarts, G. (1973). Ionospheric Processes. *Aeronomy*, 124–151. <https://doi.org/10.1016/B978-0-12-077802-7.50010-5>
- Barr, R., Jones, D. L., & Rodger, C. J. (2000). ELF and VLF radio waves. *Journal of Atmospheric and Solar-Terrestrial Physics*, 62(17–18), 1689–1718. [https://doi.org/10.1016/S1364-6826\(00\)00121-8](https://doi.org/10.1016/S1364-6826(00)00121-8)
- Basu, S., Basu, S., Rich, F. J., Groves, K. M., MacKenzie, E., Coker, C., ... Becker-Guedes, F. (2007). Response of the equatorial ionosphere at dusk to penetration electric fields during intense magnetic storms. *Journal of Geophysical Research: Space Physics*, 112(8), 1–14. <https://doi.org/10.1029/2006JA012192>
- Belrose, J. S., & Burke, M. J. (1964). Study of the lower ionosphere using partial reflection: 1. Experimental technique and method of analysis. *Journal of Geophysical Research*, 69(13), 2799. <https://doi.org/10.1029/JZ069i013p02799>
- Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning Long Term Dependencies with Gradient Descent is Difficult. *IEEE Transactions on Neural Networks*, 5(2), 157–166. <https://doi.org/10.1109/72.279181>
- Blanc, M., & Richmond, A. D. (1980). The ionospheric disturbance dynamo. *Journal of Geophysical Research*, 85(A4), 1669–1686. <https://doi.org/10.1029/JA085iA04p01669>
- Boberg, F., Wintoft, P., & Lundstedt, H. (2000). Real time Kp predictions from solar wind data using neural networks. *Phys Chem Earth*, 25(4), 275–280. [https://doi.org/10.1016/S1464-1917\(00\)00016-7](https://doi.org/10.1016/S1464-1917(00)00016-7)
- Brum, C. G. M., Rodrigues, F. da S., dos Santos, P. T., Matta, A. C., Aponte, N., Gonzalez, S. A., & Robles, E. (2011). A modeling study of foF2 and hmF2 parameters measured by the Arecibo incoherent scatter radar and comparison with IRI model predictions for solar cycles 21, 22, and 23. *Journal of Geophysical Research: Space Physics*, 116(A3), 1–12. <https://doi.org/10.1029/2010JA015727>
- Budden, K. G. (1985). *The propagation of radio waves*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511564321>
- Buonsanto, M. J. (1999). Ionospheric storms – A review. *Space Science Reviews*, 88(3–4), 563–601. <https://doi.org/10.1023/A:1005107532631>
- Burden, F., & Winkler, D. (2009). Bayesian Regularization of Neural Networks. In D. J. Livingstone (Ed.), *Artificial Neural Networks: Methods and Applications* (pp. 23–42).

- Totowa, NJ: Humana Press. https://doi.org/10.1007/978-1-60327-101-1_3
- Burešová, D., & Laštovička, J. (2007). Pre-storm enhancements of foF2 above Europe. *Advances in Space Research*, 39(8), 1298–1303. <https://doi.org/10.1016/j.asr.2007.03.003>
- Chapman, S., & Cowling, T. G. (1970). *The Mathematical Theory of Non-uniform Gases: An Account of the Kinetic Theory of Viscosity, Thermal Conduction and Diffusion in Gases*. Cambridge University Press. Retrieved from <https://books.google.co.jp/books?id=Cbp5JP2OTrwC>
- Chow, T., & Cho, S.-Y. (2007). *Neural Networks and Computing Learning Algorithms and Applications*. Imperial College Press. Retrieved from <https://books.google.co.jp/books?id=q9ou0N5qBBMC>
- Chow, T. W. S., & Fang, Y. (1998). A recurrent neural-network-based real-time learning control strategy applying to nonlinear systems with unknown dynamics. *IEEE Transactions on Industrial Electronics*, 45(1), 151–161. <https://doi.org/10.1109/41.661316>
- Clark, R. ., Yeh, K. ., & Liu, C. . (1971). Interaction of internal gravity waves with the ionospheric F2-layer. *Journal of Atmospheric and Terrestrial Physics*, 33(10), 1567–1576. [https://doi.org/10.1016/0021-9169\(71\)90074-2](https://doi.org/10.1016/0021-9169(71)90074-2)
- Clemesha, B., Takahashi, H., Simonich, D., Gobbi, D., & Batista, P. (2005). Experimental evidence for solar cycle and long-term change in the low-latitude MLT region. *Journal of Atmospheric and Solar-Terrestrial Physics*, 67(1–2), 191–196. <https://doi.org/10.1016/j.jastp.2004.07.027>
- Clilverd, M. A., Thomson, N. R., & Rodger, C. J. (1999). Sunrise effects on VLF signals propagating over a long north-south path. *Radio Science*, 34(4), 939–948. <https://doi.org/10.1029/1999RS900052>
- Clilverd, M. A., Rodger, C. J., Millan, R. M., Sample, J. G., Kokorowski, M., McCarthy, M. P., ... Spanswick, E. (2007). Energetic particle precipitation into the middle atmosphere triggered by a coronal mass ejection. *Journal of Geophysical Research: Space Physics*, 112(12), 1–12. <https://doi.org/10.1029/2007JA012395>
- Clilverd, M. a., Rodger, C. J., Thomson, N. R., Brundell, J. B., Ulich, T., Lichtenberger, J., ... Turunen, E. (2009). Remote sensing space weather events: Antarctic-Arctic Radiation-belt (Dynamic) Deposition-VLF Atmospheric Research Konsortium network. *Space Weather*, 7(4). <https://doi.org/10.1029/2008SW000412>
- Clilverd, M. a, Rodger, C. J., Thomson, N. R., Lichtenberger, J., Steinbach, P., Cannon, P., &

- Angling, M. J. (2001). Total solar eclipse effects on VLF signals: Observations and modeling. *Radio Science*, 36(4), 773–788. <https://doi.org/10.1029/2000RS002395>
- Cohen, M. B., Inan, U. S., & Fishman, G. (2006). Terrestrial gamma ray flashes observed aboard the Compton Gamma Ray Observatory/Burst and Transient Source Experiment and ELF/VLF radio atmospherics. *Journal of Geophysical Research: Atmospheres*, 111(24). <https://doi.org/10.1029/2005JD006987>
- Cullen, K. E., Rey, C. G., Guitton, D., & Galiana, H. L. (1996). The use of system identification techniques in the analysis of oculomotor burst neuron spike train dynamics. *Journal of Computational Neuroscience*, 3(4), 347–368. <https://doi.org/10.1007/BF00161093>
- Cummer, S. A., Bell, T. F., Inan, U. S., & Chenette, D. L. (1997). VLF remote sensing of high-energy auroral particle precipitation. *Journal of Geophysical Research: Space Physics*, 102(A4), 7477–7484. <https://doi.org/10.1029/96JA03721>
- Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems*, 2(4), 303–314. <https://doi.org/10.1007/BF02551274>
- Dan Foresee, F., & Hagan, M. T. (1997). Gauss-Newton approximation to Bayesian learning. In *Proceedings of International Conference on Neural Networks (ICNN'97)* (Vol. 3, pp. 1930–1935). IEEE. <https://doi.org/10.1109/ICNN.1997.614194>
- Danilov, A. D. (1998). Solar activity effects in the ionospheric D region. *Annales Geophysicae*, 16(12), 1527. <https://doi.org/10.1007/s005850050719>
- Davies, K. (1989). *Ionospheric Radio*. Peter Peregrinus Ltd., London, United Kingdom.
- Deyfus, G. (2005). *Neural Networks*. Berlin/Heidelberg: Springer-Verlag. <https://doi.org/10.1007/3-540-28847-3>
- Dieminger, W., Hartmann, G. K., & Leitinger, R. (1996). *The upper atmosphere: data analysis and interpretation*. Springer-Verlag. Retrieved from <https://books.google.co.jp/books?id=D1sRAQAIAAJ>
- DiPietro, R., Rupprecht, C., Navab, N., & Hager, G. D. (2017). Analyzing and Exploiting NARX Recurrent Neural Networks for Long-Term Dependencies. Retrieved from <http://arxiv.org/abs/1702.07805>
- Dowden, R. L., & Adams, C. D. D. (1988). Phase and amplitude perturbations on subionospheric signals explained in terms of echoes from lightning-induced electron precipitation ionization patches. *Journal of Geophysical Research: Space Physics*, 93(A10), 11543–11550. <https://doi.org/10.1029/JA093iA10p11543>
- Dowden, R. L., & Adams, C. D. D. (2006). softPAL. In *2nd VERSIM Workshop 2006 at Sodankylä Geophysical Observatory (SGO)* (p. 1).

- Duch, W., & Jankowski, N. (1999). Survey of neural transfer functions. *Neural Computing Surveys*, 2, 163–212. Retrieved from ftp://ftp.icsi.berkeley.edu/pub/ai/jagota/vol2_6.pdf
- Engelbrecht, A. P. (2007). *Computational Intelligence*. Chichester, UK: John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470512517>
- Evans, J. V. (1969). Theory and practice of ionosphere study by Thompson scatter radar. *Proc. IEEE*, 57(4), 496–530.
- Fejer, B. G., Blanc, M., & Richmond, A. D. (2017). Post-Storm Middle and Low-Latitude Ionospheric Electric Fields Effects. *Space Science Reviews*, 206(1–4), 407–429. <https://doi.org/10.1007/s11214-016-0320-x>
- Ferreira, A. A., Borges, R. A., Paparini, C., Ciraolo, L., & Radicella, S. M. (2017). Short-term estimation of GNSS TEC using a neural network model in Brazil. *Advances in Space Research*, 60(8), 1765–1776. <https://doi.org/10.1016/j.asr.2017.06.001>
- Fesen, C. G., Roble, R. G., & Duboin, M.-L. (1995). Simulations of seasonal and geomagnetic activity effects at Saint Santin. *Journal of Geophysical Research*, 100407(1), 397–21. <https://doi.org/10.1029/95JA01211>
- Forney, G. D. (1973). The viterbi algorithm. *Proceedings of the IEEE*, 61(3), 268–278. <https://doi.org/10.1109/PROC.1973.9030>
- Garson, G. D. (1991). Interpreting Neural-network Connection Weights. *AI Expert*, 6(4), 46–51. Retrieved from <http://dl.acm.org/citation.cfm?id=129449.129452>
- Ghosh, S. N. (2002). *Electromagnetic Theory and Wave Propagation*. CRC Press. Retrieved from <https://books.google.co.jp/books?id=6Mvf4-gsVycC>
- De Giorgi, M. G., Malvoni, M., & Congedo, P. M. (2016). Comparison of strategies for multi-step ahead photovoltaic power forecasting models based on hybrid group method of data handling networks and least square support vector machine. *Energy*, 107, 360–373. <https://doi.org/10.1016/j.energy.2016.04.020>
- Giri, F., & Bai, E. W. (2010). *Block-oriented Nonlinear System Identification*. (F. Giri & E.-W. Bai, Eds.), *Lecture notes in control and information sciences* (Vol. 404). London: Springer London. <https://doi.org/10.1007/978-1-84996-513-2>
- Goh, A. T. C. (1995). Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering*, 9(3), 143–151. [https://doi.org/10.1016/0954-1810\(94\)00011-S](https://doi.org/10.1016/0954-1810(94)00011-S)
- Gronemeyer, S. A., & McBride, A. L. (1976). MSK and Offset QPSK Modulation. *IEEE Transactions on Communications*, 24(8), 809–820. <https://doi.org/10.1109/TCOM.1976.1093392>

- Hagan, M. T., Demuth, H. B., Beale, M. H., & De Jesús, O. (2014). *Neural Network Design*. Martin Hagan. Retrieved from <http://books.google.ru/books?id=bUNJAAAACAAJ>
- Haigh, J. D. (2007). The Sun and the Earth ' s Climate Imprint / Terms of Use. *Living Reviews in Solar Physics*.
- Halpern, J. Y. (2001). Conditional plausibility measures and Bayesian networks. *Journal of Artificial Intelligence Research*, 14, 369–399. <https://doi.org/10.1613/jair.817>
- Hargreaves, J. K. (1992). *The solar–terrestrial environment*. Cambridge Atmospheric and Space Science (Vol. 5). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511628924>
- Hayakawa, M. (1989). Satellite observation of low-latitude VLF radio noises and their association with thunderstorms. *Journal of Geomagnetism and Geoelectricity*, 41(7), 573–595. <https://doi.org/10.5636/jgg.41.573>
- Hayakawa, M. (2007). VLF/LF Radio Sounding of Ionospheric Perturbations Associated with Earthquakes. *Sensors*, 7(7), 1141–1158. <https://doi.org/10.3390/s7071141>
- Hayakawa, M., & Hobara, Y. (2010). Current status of seismo-electromagnetics for short-term earthquake prediction. *Geomatics, Natural Hazards and Risk*, 1(2), 115–155. <https://doi.org/10.1080/19475705.2010.486933>
- Haykin, S. (2009). *Neural Networks: A Comprehensive Foundation*. <https://doi.org/10987654321>
- Hegai, V. V., Kim, V. P., & Legen'ka, A. D. (2017). Ionospheric F2-layer Perturbations Observed After the M8.8 Chile Earthquake on February 27, 2010, at Long Distance from the Epicenter. *Journal of Astronomy and Space Sciences*, 34(1), 1–5. <https://doi.org/10.5140/JASS.2017.34.1.1>
- Hegai, V. V., & Kim, V. P. (2016). Disturbance in the Daytime Midlatitude Upper F Region Associated with a Medium Scale Electrodynamic Vortex Motion of Plasma. *Journal of Astronomy and Space Sciences*, 33(3), 207–210. <https://doi.org/10.5140/JASS.2016.33.3.207>
- Hellerstein, J. L., Diao, Y., Parekh, S., & Tilbury, D. M. (2004). *Feedback Control of Computing Systems*. Wiley.
- Helliwell, R. a., Katsufakis, J. P., & Trimpi, M. L. (1973). Whistler-induced amplitude perturbation in VLF propagation. *Journal of Geophysical Research*. <https://doi.org/10.1029/JA078i022p04679>
- Hobara, Y., Iwasaki, N., Hayashida, T., Hayakawa, M., Ohta, K., & Fukunishi, H. (2001). Interrelation between ELF transients and ionospheric disturnances in association with

- Sprites and Elves. *Geophysical Research Letters*, 28(5), 935–938. <https://doi.org/10.1029/2000GL003795>
- Hobara, Y., Molchanov, O. A., Hayakawa, M., & Ohta, K. (1995). Propagation characteristics of whistler waves in the Jovian ionosphere and magnetosphere. *Journal of Geophysical Research*, 100(A12), 23523. <https://doi.org/10.1029/95JA02434>
- Horne, R. B., Thorne, R. M., Glauert, S. A., Meredith, N. P., Pokhotelov, D., & Santolík, O. (2007). Electron acceleration in the Van Allen radiation belts by fast magnetosonic waves. *Geophysical Research Letters*, 34(17), 2–6. <https://doi.org/10.1029/2007GL030267>
- Hu, X., & Wang, J. (2006). Solving pseudomonotone variational inequalities and pseudoconvex optimization problems using the projection neural network. *IEEE Transactions on Neural Networks*, 17(6), 1487–1499. <https://doi.org/10.1109/TNN.2006.879774>
- Huang, C. S., Foster, J. C., Goncharenko, L. P., Erickson, P. J., Rideout, W., & Coster, A. J. (2005). A strong positive phase of ionospheric storms observed by the Millstone Hill incoherent scatter radar and global GPS network. *Journal of Geophysical Research: Space Physics*, 110(A6), 1–13. <https://doi.org/10.1029/2004JA010865>
- Huang, S. N., Tan, K. K., & Lee, T. H. (2005). Further result on a dynamic recurrent neural-network-based adaptive observer for a class of nonlinear systems. *Automatica*, 41(12), 2161–2162. <https://doi.org/10.1016/j.automatica.2005.07.003>
- Hunsucker, R. D., & Hargreaves, J. K. (2002). *The high-latitude ionosphere and its effects on radio propagation*. (A. J. Dessler, J. T. Houghton, & M. J. Rycroft, Eds.). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511535758>
- Ikubanni, S. O., & Adeniyi, J. O. (2012). On the Dependence of F2 layer critical frequency on F10.7 solar flux at an Equatorial station. *World Journal of Engineering and Pure and Applied Science*, 3(2), 85–91. Retrieved from www.rrpjournals.com
- Ikubanni, S. O., Adebisin, B. O., Adebisi, S. J., & Adeniyi, J. O. (2013). Relationship between F2 layer critical frequency and solar activity indices during different solar epochs. *Indian Journal of Radio and Space Physics*, 42(2), 73–81.
- Inan, U. S., & Carpenter, D. L. (1987). Lightning-induced electron precipitation events observed at L ~ 2.4 as phase and amplitude perturbations on subionospheric VLF signals. *Journal of Geophysical Research*, 92(A4), 3293. <https://doi.org/10.1029/JA092iA04p03293>
- Inan, U. S., Reising, S. C., Fishman, G. J., & Horack, J. M. (1996). On the association of terrestrial gamma-ray bursts with lightning and implications for sprites. *Geophysical*

- Research Letters*, 23(9), 1017–1020. <https://doi.org/10.1029/96GL00746>
- Inan, U. S., Cummer, S. A., & Marshall, R. A. (2010). A survey of ELF and VLF research on lightning-ionosphere interactions and causative discharges. *Journal of Geophysical Research: Space Physics*, 115(6), 1–21. <https://doi.org/10.1029/2009JA014775>
- Inan, U. S., Inan, A., & Said, R. (2015). *Engineering Electromagnetics and Waves*. Pearson Education.
- Inui, D., & Hobara, Y. (2014). Spatio-temporal characteristics of sub-ionospheric perturbations associated with annular solar eclipse over Japan: Network observations and modeling. In *2014 XXXIth URSI General Assembly and Scientific Symposium (URSI GASS)* (pp. 1–3). IEEE. <https://doi.org/10.1109/URSIGASS.2014.6929555>
- Jain, V. K., & Singh, B. (1990). Storm-time precipitation of resonant electrons at the lower edge of the inner radiation belt. *Planetary and Space Science*, 38(6), 785–790. [https://doi.org/10.1016/0032-0633\(90\)90037-Q](https://doi.org/10.1016/0032-0633(90)90037-Q)
- Jakowski, N., Wehrenpfennig, A., Heise, S., & Noack, T. (2001). Space Weather Effects in the Ionosphere and Their Impact on Positioning. *Space Weather Workshop*, 17–19.
- Kalikhman, A. D. (1980). Medium-scale travelling ionospheric disturbances and thermospheric winds in the F-region. *Journal of Atmospheric and Terrestrial Physics*, 42(8), 697–703. [https://doi.org/10.1016/0021-9169\(80\)90053-7](https://doi.org/10.1016/0021-9169(80)90053-7)
- Kiş, Ö., & Uncuoğlu, E. (2005). Comparison of three back-propagation training algorithms for two case studies. *Indian Journal of Engineering and Materials Sciences*, 12(5), 434–442. Retrieved from <http://hdl.handle.net/123456789/8460>
- Kleimenova, N., Kozyreva, O., Rozhnoy, A. A., & Solovieva, M. S. (2004). Variations in the VLF signal parameters on the Australia-Kamchatka radio path during magnetic storms. *Geomag. Aeron.*, 44(September 2017), 385–393.
- Klimenko, M. V, Klimenko, V. V, Zakharenkova, I. E., Ratovsky, K. G., Korenkova, N. A., Yasyukevich, Y. V., ... Cherniak, I. V. (2017). Similarity and differences in morphology and mechanisms of the foF2 and TEC disturbances during the geomagnetic storms on 26–30 September 2011. *Annales Geophysicae*, 35(4), 923–938. <https://doi.org/10.5194/angeo-35-923-2017>
- Kolarski, A., & Grubor, D. (2014). Sensing the Earth's low ionosphere during solar flares using VLF signals and goes solar X-ray data. *Advances in Space Research*, 53(11), 1595–1602. <https://doi.org/10.1016/j.asr.2014.02.022>
- Kumar, A., & Kumar, S. (2014). Space weather effects on the low latitude D-region ionosphere during solar minimum. *Earth, Planets and Space*, 66(1), 76.

5981-66-76

- Kumar, K. V., Maurya, A. K., Kumar, S., & Singh, R. (2016). 22 July 2009 total solar eclipse induced gravity waves in ionosphere as inferred from GPS observations over EIA. *Advances in Space Research*, 58(9), 1755–1762. <https://doi.org/10.1016/j.asr.2016.07.019>
- Kumar, S., Kumar, A., Menk, F., Maurya, A. K., Singh, R., & Veenadhari, B. (2015). Response of the low-latitude D region ionosphere to extreme space weather event of 14-16 December 2006. *Journal of Geophysical Research: Space Physics*, 120(1), 788–799. <https://doi.org/10.1002/2014JA020751>
- Kunitsyn, V. E., & Tereshchenko, E. D. (2003). *Ionospheric Tomography*. Springer Berlin Heidelberg. Retrieved from <https://books.google.co.jp/books?id=ShFjV2-ARIQC>
- De La Rosa, E., Yu, W., & Li, X. (2017). Nonlinear system modeling with deep neural networks and autoencoders algorithm. *2016 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2016 - Conference Proceedings*, 2157–2162. <https://doi.org/10.1109/SMC.2016.7844558>
- Lastovicka, J. (2002). Monitoring and forecasting of ionospheric space weather—effects of geomagnetic storms. *Journal of Atmospheric and Solar-Terrestrial Physics*, 64(5–6), 697–705. [https://doi.org/10.1016/S1364-6826\(02\)00031-7](https://doi.org/10.1016/S1364-6826(02)00031-7)
- Laštovička, J., Akmaev, R. a., Beig, G., Bremer, J., Emmert, J. T., Jacobi, C., ... Ulich, T. (2008). Emerging pattern of global change in the upper atmosphere and ionosphere. *Annales Geophysicae*, 26, 1255–1268. <https://doi.org/10.5194/angeo-26-1255-2008>
- Li, G., & Shi, J. (2012). Applications of Bayesian methods in wind energy conversion systems. *Renewable Energy*, 43, 1–8. <https://doi.org/10.1016/j.renene.2011.12.006>
- Lin, T., Horne, B. G., & Giles, C. L. (1998). How embedded memory in recurrent neural network architectures helps learning long-term temporal dependencies. *Neural Networks*, 11(5), 861–868. [https://doi.org/10.1016/S0893-6080\(98\)00018-5](https://doi.org/10.1016/S0893-6080(98)00018-5)
- Ljung, L. (1999). *System Identification: Theory for the User*. Prentice Hall PTR. Retrieved from <https://books.google.co.jp/books?id=nHFoQgAACAAJ>
- Lobzin, V. V., & Pavlov, A. V. (2002). G condition in the F2 region peak electron density: a statistical study. *Annales Geophysicae*, 20(4), 523–537. Retrieved from <http://www.ann-geophys.net/20/523/2002/>
- Lundstedt, H. (2002). Operational forecasts of the geomagnetic Dst index. *Geophysical Research Letters*, 29(24), 1–4. <https://doi.org/10.1029/2002GL016151>
- MacKay, D. J. C. (1992a). A Practical Bayesian Framework for Backprop Networks. *Neural*

- Computation*, 4(3), 448–472. <https://doi.org/10.1162/neco.1992.4.3.448>
- MacKay, D. J. C. (1992b). Bayesian Interpolation. *Neural Computation*, 4(3), 415–447. <https://doi.org/10.1162/neco.1992.4.3.415>
- Mandic, D. P., & Chambers, J. A. (2001). *Recurrent Neural Networks for Prediction* (Vol. 4). Chichester, UK: John Wiley & Sons, Ltd. <https://doi.org/10.1002/047084535X>
- Matsushita, S., & Campbell, W. H. (1967). *Physics of geomagnetic phenomena: International Geophysics series, volume I*. Academic Press INC. London LTD. <https://doi.org/10.1002/qj.49709440127>
- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4), 115–133. <https://doi.org/10.1007/BF02478259>
- McNamara, L. F. (1991). *The Ionosphere: Communications, Surveillance, and Direction Finding*. Krieger Publishing Company. Retrieved from <https://books.google.co.jp/books?id=1a1ZAAAAYAAJ>
- Mechtly, E. A., Bowhill, S. A., Smith, L. G., & Knoebel, H. W. (1967). Lower ionosphere electron concentration and collision frequency from rocket measurements of Faraday rotation, differential absorption, and probe current. *Journal of Geophysical Research*, 72(21), 5239–5245. <https://doi.org/10.1029/JZ072i021p05239>
- Moler, W. F. (1960). VLF propagation effects of a D -region layer produced by cosmic rays. *Journal of Geophysical Research*, 65(5), 1459–1468. <https://doi.org/10.1029/JZ065i005p01459>
- Moody, J. E., Hanson, S. J., & Moody, J. E. (1992). An Analysis of Generalization and Regularization in Nonlinear Learning Systems. *Neural Information Processing Systems (NIPS)*, (1), 847–854.
- Moody, J. E. (1991). Note on generalization, regularization and architecture selection\nin nonlinear learning systems. *Neural Networks for Signal Processing Proceedings of the 1991 IEEE Workshop*, 1–10. <https://doi.org/10.1109/NNSP.1991.239541>
- Moore, R. C., Barrington-Leigh, C. P., Inan, U. S., & Bell, T. F. (2003). Early/fast VLF events produced by electron density changes associated with sprite halos. *Journal of Geophysical Research: Space Physics*, 108(A10), 1–8. <https://doi.org/10.1029/2002JA009816>
- Nakamura, M. I., Maruyama, T., & Shidama, Y. (2007). Using a neural network to make operational forecasts of ionospheric variations and storms at Kokubunji, Japan. *Earth, Planets and Space*, 59(12), 1231–1239. <https://doi.org/10.1186/BF03352071>
- Nakamura, M., Maruyama, T., & Shidama, Y. (2009). Using a neural network to make

- operational forecasts of ionospheric variations and storms at Kokubunji, Japan. *Journal of the National Institute of Information and Communications Technology*, 56(1–4), 391–406.
- Ogunfunmi, T. (2007). *Adaptive Nonlinear System Identification: The Volterra and Weiner Series Approaches*. Springer Santa Clara.
- Pakhomov, S. V., & Gorbunov, A. N. (1983). ne(h) profiles of the ionospheric D region of the equatorial zone measured in the period of solar activity maximum. *Geomagnetism and Aeronomy*, 23, 134–136.
- Pal, S., & Hobara, Y. (2016). Mid-latitude atmosphere and ionosphere connection as revealed by very low frequency signals. *Journal of Atmospheric and Solar-Terrestrial Physics*, 138–139, 227–232. <https://doi.org/10.1016/j.jastp.2015.12.008>
- Palit, A. K., & Popovic, D. (2006). *Computational Intelligence in Time Series Forecasting: Theory and Engineering Applications*. Springer London. Retrieved from <http://amstat.tandfonline.com/doi/abs/10.1198/tech.2006.s425>
- Pan, X., Lee, B., & Zhang, C. (2013). A comparison of neural network backpropagation algorithms for electricity load forecasting. *Proceedings - 2013 IEEE International Workshop on Intelligent Energy Systems, IWIES 2013*, 22–27. <https://doi.org/10.1109/IWIES.2013.6698556>
- Pedroni, N., Zio, E., & Apostolakis, G. E. (2010). Comparison of bootstrapped artificial neural networks and quadratic response surfaces for the estimation of the functional failure probability of a thermal-hydraulic passive system. *Reliability Engineering and System Safety*, 95(4), 386–395. <https://doi.org/10.1016/j.ress.2009.11.009>
- Peter, W. B., Chevalier, M. W., & Inan, U. S. (2006). Perturbations of midlatitude subionospheric VLF signals associated with lower ionospheric disturbances during major geomagnetic storms. *Journal of Geophysical Research: Space Physics*, 111(3), 1–14. <https://doi.org/10.1029/2005JA011346>
- Prechelt, L. (1998). Automatic early stopping using cross validation: Quantifying the criteria. *Neural Networks*, 11(4), 761–767. [https://doi.org/10.1016/S0893-6080\(98\)00010-0](https://doi.org/10.1016/S0893-6080(98)00010-0)
- Prolss, G. W. (1995). Ionospheric F-Region Storms. In *Handbook of Atmospheric Electrodynamics*. CRC Press.
- Reid, G. C. (1972). Ionospheric Effects of Solar Activity. In *Solar Activity Observations and Predictions* (pp. 293–312). New York: American Institute of Aeronautics and Astronautics. <https://doi.org/10.2514/5.9781600865046.0293.0312>
- Reinisch, B. W., Haines, D. M., Bibl, K., Galkin, I., Huang, X., Kitrosser, D. F., ... Scali, J. L.

- (1997). Ionospheric sounding in support of over-the-horizon radar. *Radio Science*, 32(4), 1681–1694. <https://doi.org/10.1029/97RS00841>
- Revallo, M., Valach, F., Hejda, P., & Bochníček, J. (2014). A neural network Dst index model driven by input time histories of the solar wind-magnetosphere interaction. *Journal of Atmospheric and Solar-Terrestrial Physics*, 110–111(January), 9–14. <https://doi.org/10.1016/j.jastp.2014.01.011>
- Rieger, M., & Leitingner, R. (2002). Assessment of TID activity from GPS phase data collected in a dense network of GPS receivers. *Acta Geodaetica et Geophysica Hungarica*, 37(2), 327–341. <https://doi.org/10.1556/AGeod.37.2002.2-3.23>
- Rishbeth, H., & Garriott, O. K. (1969). *Introduction to ionospheric physics*. Academic Press. Retrieved from <https://books.google.co.jp/books?id=YdbAAAAIAAJ>
- Rodger, C. J. (1999). Red sprites, upward lightning, and VLF perturbations. *Reviews of Geophysics*, 37(3), 317. <https://doi.org/10.1029/1999RG900006>
- Rozhnoi, A., Solovieva, M., Hayakawa, M., Yamaguchi, H., Hobara, Y., Levin, B., & Fedun, V. (2014). Tsunami-driven ionospheric perturbations associated with the 2011 Tohoku earthquake as detected by subionospheric VLF signals. *Geomatics, Natural Hazards and Risk*, 5(4), 285–292. <https://doi.org/10.1080/19475705.2014.888100>
- Rozhnoi, a., Solovieva, M., Parrot, M., Hayakawa, M., Biagi, P.-F., Schwingenschuh, K., & Fedun, V. (2015). VLF/LF signal studies of the ionospheric response to strong seismic activity in the Far Eastern region combining the DEMETER and ground-based observations. *Physics and Chemistry of the Earth, Parts A/B/C*. <https://doi.org/10.1016/j.pce.2015.02.005>
- Sanghai, S., Domingos, P., & Weld, D. (2005). Relational dynamic bayesian networks. *Journal of Artificial Intelligence Research*, 24, 759–797.
- Schmitter, E. D. (2011). Remote sensing planetary waves in the midlatitude mesosphere using low frequency transmitter signals. *Annales Geophysicae*, 29(7), 1287–1293. <https://doi.org/10.5194/angeo-29-1287-2011>
- Schmitter, E. D. (2013). Modeling solar flare induced lower ionosphere changes using VLF/LF transmitter amplitude and phase observations at a midlatitude site. *Annales Geophysicae*, 31(4), 765–773. <https://doi.org/10.5194/angeo-31-765-2013>
- Schunk, R., & Nagy, A. (2009). *Ionospheres. Eos, Transactions American Geophysical Union* (Vol. 82). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511635342>
- Senior, C., & Blanc, M. (1984). On the Control of Magnetospheric Convection by the Spatial

- Distribution of Ionospheric Conductivities. *Journal of Geophysical Research*, 89(A1), 261–284. <https://doi.org/10.1029/JA089iA01p00261>
- Silber, I., Price, C., Rodger, C. J., & Haldoupis, C. (2013). Links between mesopause temperatures and ground-based VLF narrowband radio signals. *Journal of Geophysical Research: Atmospheres*, 118(10), 4244–4255. <https://doi.org/10.1002/jgrd.50379>
- Silber, I., Price, C., & Rodger, C. J. (2016). Semi-annual oscillation (SAO) of the nighttime ionospheric D region as detected through ground-based VLF receivers. *Atmospheric Chemistry and Physics*, 16(5), 3279–3288. <https://doi.org/10.5194/acp-16-3279-2016>
- Singh, B., Tyagi, R., Hobara, Y., & Hayakawa, M. (2014). X-rays and solar proton event induced changes in the first mode Schumann resonance frequency observed at a low latitude station Agra, India. *Journal of Atmospheric and Solar-Terrestrial Physics*, 113(2014), 1–9. <https://doi.org/10.1016/j.jastp.2014.02.010>
- Stubbe, P. (1975). The effect of neutral winds on the seasonal F-region variation. *Journal of Atmospheric and Terrestrial Physics*, 37(4), 675–679. [https://doi.org/10.1016/0021-9169\(75\)90063-X](https://doi.org/10.1016/0021-9169(75)90063-X)
- Tatsuta, K., Hobara, Y., Pal, S., & Balikhin, M. (2015). Sub-ionospheric VLF signal anomaly due to geomagnetic storms: a statistical study. *Annales Geophysicae*, 33(11), 1457–1467. <https://doi.org/10.5194/angeo-33-1457-2015>
- Thomson, N. R., Rodger, C. J., & Clilverd, M. A. (2005). Large solar flares and their ionospheric D region enhancements. *Journal of Geophysical Research: Space Physics*, 110(A6), 1–10. <https://doi.org/10.1029/2005JA011008>
- Thomson, N. R., Clilverd, M. a., & McRae, W. M. (2007). Nighttime ionospheric D region parameters from VLF phase and amplitude. *Journal of Geophysical Research: Space Physics*, 112(7), 1–14. <https://doi.org/10.1029/2007JA012271>
- Thomson, N. R., Clilverd, M. A., & Rodger, C. J. (2014). Low-latitude ionospheric D region dependence on solar zenith angle. *Journal of Geophysical Research: Space Physics*, 119(8), 6865–6875. <https://doi.org/10.1002/2014JA020299>
- Tomko, A. A., & Hepner, T. (2001). Worldwide monitoring of VLF-LF propagation and atmospheric noise. *Radio Science*, 36(2), 363–369. <https://doi.org/10.1029/1999RS002402>
- Trísková, L. (1994). The nighttime mid-latitude foF2 and geomagnetic indices. *Advances in Space Research*, 14(12), 153–156. [https://doi.org/10.1016/0273-1177\(94\)90258-5](https://doi.org/10.1016/0273-1177(94)90258-5)
- Tsagouri, I., Belehaki, A., Moraitis, G., & Mavromichalaki, H. (2000). Positive and negative ionospheric disturbances at middle latitudes during geomagnetic storms. *Geophysical*

- Research Letters*, 27(21), 3579–3582. <https://doi.org/10.1029/2000GL003743>
- Tsoi, A. C., & Back, A. (1997). Discrete time recurrent neural network architectures: A unifying review. *Neurocomputing*, 15(3–4), 183–223. [https://doi.org/10.1016/S0925-2312\(97\)00161-6](https://doi.org/10.1016/S0925-2312(97)00161-6)
- Tulunay, E., Senalp, E. T., Radicella, S. M., & Tulunay, Y. (2006). Forecasting total electron content maps by neural network technique. *Radio Science*, 41(4), 1–12. <https://doi.org/10.1029/2005RS003285>
- Villain, J. P., Greenwald, R. A., & Vickrey, J. F. (1984). HF ray tracing at high latitudes using measured meridional electron density distributions. *Radio Science*, 19(1), 359–374. <https://doi.org/10.1029/RS019i001p00359>
- Volland, H., & Warnecke, G. (1968). Stratospheric-Mesospheric Coupling During Stratospheric Warmings. In Intergovernmental Panel on Climate Change (Ed.), *Climate Change 2013 - The Physical Science Basis* (pp. 1–48). Maryland: NASA. <https://doi.org/NASA-TM-X-63376,X-621-68-369>
- Wait, J. (1957). The mode theory of VLF ionospheric propagation for finite ground conductivity. *Proceedings of the IRE*, 1956, 760–767. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4056596
- Wang, R., Zhou, C., Deng, Z., Ni, B., & Zhao, Z. (2013). Predicting foF2 in the China region using the neural networks improved by the genetic algorithm. *Journal of Atmospheric and Solar-Terrestrial Physics*, 92(2013), 7–17. <https://doi.org/10.1016/j.jastp.2012.09.010>
- Wasserman, P. D. (1989). *Neural Computing: Theory and Practice*. New York, NY, USA: Van Nostrand Reinhold Co.
- Webber, W. (1962). The production of free electrons in the ionospheric D Layer by solar and galactic cosmic rays and the resultant absorption of radio waves. *Journal of Geophysical Research*, 67(13), 5091–5106. <https://doi.org/10.1029/JZ067i013p05091>
- Wichaipanich, N., Hozumi, K., Supnithi, P., & Tsugawa, T. (2017). A comparison of neural network-based predictions of foF2 with the IRI-2012 model at conjugate points in Southeast Asia. *Advances in Space Research*, 59(12), 2934–2950. <https://doi.org/10.1016/j.asr.2017.03.023>
- Wolf, R. (1856). Mitteilungen uber die Sonnenflecken. *Astron. Mitt. Zurich*, 1(8).
- Yang, Y. F., & Chen, C. X. (2009). Influence of seasonal variation of neutral winds on the coupling between North and South Hemispheres in a realistic Earth main field. *Science in China, Series D: Earth Sciences*, 52(10), 1661–1672. <https://doi.org/10.1007/s11430-009-0128-6>

- Yu, H., & Wilamowski, B. M. (2011). Levenberg-Marquardt training. *Industrial Electronics Handbook, Vol. 5 – Intelligent Systems*, 12-1-18. <https://doi.org/doi:10.1201/b10604-15>
- Yue, Z., Songzheng, Z., & Tianshi, L. (2011). Bayesian regularization BP Neural Network model for predicting oil-gas drilling cost. *BMEI 2011 - Proceedings 2011 International Conference on Business Management and Electronic Information*, 2, 483–487. <https://doi.org/10.1109/ICBMEI.2011.5917952>
- Zolesi, B., & Cander, L. R. (2014). The General Structure of the Ionosphere. In *Ionospheric Prediction and Forecasting* (pp. 11–48). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-38430-1_2